

Artificial intelligence (AI) in healthcare and research



OVERVIEW

- AI is being used or trialled for a range of healthcare and research purposes, including detection of disease, management of chronic conditions, delivery of health services, and drug discovery.
- AI has the potential to help address important health challenges, but might be limited by the quality of available health data, and by the inability of AI to display some human characteristics.
- The use of AI raises ethical issues, including: the potential for AI to make erroneous decisions; the question of who is responsible when AI is used to support decision-making; difficulties in validating the outputs of AI systems; inherent biases in the data used to train AI systems; ensuring the protection of potentially sensitive data; securing public trust in the development and use of AI technologies; effects on people's sense of dignity and social isolation in care situations; effects on the roles and skill-requirements of healthcare professionals; and the potential for AI to be used for malicious purposes.
- A key challenge will be ensuring that AI is developed and used in a way that is transparent and compatible with the public interest, whilst stimulating and driving innovation in the sector.

WHAT IS AI?

There is no universally agreed definition of AI. The term broadly refers to computing technologies that resemble processes associated with human intelligence, such as reasoning, learning

and adaptation, sensory understanding, and interaction.¹ Currently, most applications of AI are narrow, in that they are only able to carry out specific tasks or solve pre-defined problems.²

AI works in a range of ways, drawing on principles and tools, including from maths, logic, and biology. An important feature of contemporary AI technologies is that they are increasingly able to make sense of varied and unstructured kinds of data, such as natural language text and images. Machine-learning has

been the most successful type of AI in recent years, and is the underlying approach of many of the applications currently in use.³ Rather than following pre-programmed instructions, machine-learning allows systems to discover patterns and derive its own rules when it is presented with data and new experiences.⁴

RECENT INTEREST IN AI

AI is not new, but there have been rapid advances in the field in recent years. This has in part been enabled by developments in computing power and the huge volumes of digital data that are now generated.⁵ A wide range of applications of AI are now being explored with considerable public and private investment and interest. The UK Government announced its ambition to make the UK a world leader in AI and data technologies in its 2017 Industrial Strategy. In April 2018, a £1bn AI sector deal between UK Government and industry was announced, including £300 million towards AI research.⁶

AI is lauded as having the potential to help address important health challenges, such as meeting the care needs of an ageing population.

Major technology companies - including Google, Microsoft, and IBM - are investing in the development of AI for healthcare and research. The number of AI start-up companies has also been steadily increasing.⁷ There are several UK-based companies, some of which have been set up in collaboration with UK universities and hospitals. Partnerships have been formed between NHS providers and AI developers such as IBM, DeepMind, Babylon Health, and Ultromics.

Such partnerships have attracted controversy, and wider concerns about AI have been the focus of several inquiries and initiatives within industry, and medical and policy communities (see Box 1).

BOX 1. EXAMPLES OF INQUIRIES AND INITIATIVES ON AI

- **UK Government Centre for Data Ethics and Innovation** – announced in January 2018 to advise on safe, ethical, and innovative uses of data-driven technologies.⁸
- **Ada Lovelace Institute** – the Nuffield Foundation announced it will set up the Institute by the end of 2018 to examine ethical and social issues arising from the use of data, algorithms, and AI, ensuring they are harnessed for social well-being.⁹
- **Partnership on AI** – a platform for discussion and engagement around AI founded by Amazon, Apple, DeepMind, Facebook, Google, IBM, and Microsoft.¹⁰
- **IEEE** – launched a Global Initiative on Ethics of Autonomous and Intelligent Systems in 2016.¹¹
- **United Nations Interregional Crime and Justice Research Institute** – set up a programme on Artificial Intelligence and Robotics in 2015.¹²
- **Asilomar AI Principles** – developed in 2017 by the Future of Life Institute (US) to guide AI research and application, and signed by over 3,800 researchers and others working in AI and robotics around the world.¹³
- Reports on AI have been published by the **House of Lords Select Committee on Artificial Intelligence**,⁵ the **Royal Society**,³ **Reform**,¹⁴ **Future Advocacy and Wellcome**,¹⁵ **Nesta**,¹⁶ and the **European Group on Ethics in Science and New Technologies**.¹⁷ A further report is expected from the **House of Commons Science and Technology Select Committee**.¹⁸

APPLICATIONS OF AI IN HEALTHCARE AND RESEARCH

HEALTHCARE ORGANISATION

AI has the potential to be used in planning and resource allocation in health and social care services. For example, the IBM Watson Care Manager system is being piloted by Harrow Council with the aim of improving cost efficiency. It matches individuals with a care provider that meets their needs, within their allocated care budget. It also designs individual care plans, and claims to offer insights for more effective use of care management resources.¹⁹

AI is also being used with the aim of improving patient experience. Alder Hey Children's Hospital in Liverpool is working with IBM Watson to create a 'cognitive hospital', which will include an app to facilitate interactions with patients. The app aims to identify patient anxieties before a visit, provide information on demand, and equip clinicians with information to help them to deliver appropriate treatments.²⁰

MEDICAL RESEARCH

AI can be used to analyse and identify patterns in large and complex datasets faster and more precisely than has previously been possible.²¹ It can also be used to search the scientific literature for relevant studies, and to combine different kinds of data; for example, to aid drug discovery.²² The Institute of Cancer Research's canSAR database combines genetic and clinical data from patients with information from scientific research, and uses AI to make predictions about new targets for cancer drugs.²³ Researchers have developed an AI 'robot scientist' called Eve which is designed to make the process of drug discovery faster and more economical.²⁴ AI systems used in healthcare could also be valuable for medical research by helping to match suitable patients to clinical studies.²⁵

CLINICAL CARE

AI has the potential to aid the diagnosis of disease and is currently being trialled for this purpose in some UK hospitals. Using AI to analyse clinical data, research publications, and professional guidelines could also help to inform decisions about treatment.²⁶

Possible uses of AI in clinical care include:

- **Medical imaging** – medical scans have been systematically collected and stored for some time and are readily available to train AI systems.²⁷ AI could reduce the cost and time involved in analysing scans, potentially allowing more scans to be taken to better target treatment.⁵ AI has shown promising results in detecting conditions such as pneumonia, breast and skin cancers, and eye diseases.²⁸
- **Echocardiography** – the Ultromics system, trialled at John Radcliffe Hospital in Oxford, uses AI to analyse echocardiography scans that detect patterns of heartbeats and diagnose coronary heart disease.²⁹
- **Screening for neurological conditions** – AI tools are being developed that analyse speech patterns to predict psychotic episodes and identify and monitor symptoms of neurological conditions such as Parkinson's disease.³⁰
- **Surgery** – robotic tools controlled by AI have been used in research to carry out specific tasks in keyhole surgery, such as tying knots to close wounds.³¹

PATIENT AND CONSUMER-FACING APPLICATIONS

Several apps that use AI to offer personalised health assessments and home care advice are currently on the market. The app Ada Health Companion uses AI to operate a chat-bot, which combines information about symptoms from the user with other information to offer possible diagnoses.³² GP at Hand, a similar app developed by Babylon Health, is currently being trialled by a group of NHS surgeries in London.³³

Information tools or chat-bots driven by AI are being used to help with the management of chronic medical conditions. For example, the Arthritis Virtual Assistant developed by IBM for Arthritis Research UK is learning through interactions with patients to provide personalised information and advice concerning medicines, diet, and exercise.³⁴ Government-funded and commercial initiatives are exploring ways in which AI could be used to power robotic systems and apps to support people living at home with conditions such as early stage dementia,

potentially reducing demands on human care workers and family carers.³⁵

AI apps that monitor and support patient adherence to prescribed medication and treatment have been trialled with promising results, for example, in patients with tuberculosis.³⁶ Other tools, such as Sention, use AI to analyse information collected by sensors worn by patients at home. The aim is to detect signs of deterioration to enable early intervention

and prevent hospital admissions.³⁷

PUBLIC HEALTH

AI has the potential to be used to aid early detection of infectious disease outbreaks and sources of epidemics, such as water contamination.³⁸ AI has also been used to predict adverse drug reactions, which are estimated to cause up to 6.5 per cent of hospital admissions in the UK.³⁹

LIMITS OF AI

AI depends on digital data, so inconsistencies in the availability and quality of data restrict the potential of AI. Also, significant computing power is required for the analysis of large and complex data sets. While many are enthusiastic about the possible uses of AI in the NHS, others point to the practical challenges, such as the fact that medical records are not consistently digitised across the NHS, and the lack of interoperability and standardisation in NHS IT systems, digital record keeping, and data labelling.⁵ There are questions about the extent to which patients and doctors are comfortable with digital sharing of

personal health data.⁴⁰

Humans have attributes that AI systems might not be able to authentically possess, such as compassion.⁴¹ Clinical practice often involves complex judgments and abilities that AI currently is unable to replicate, such as contextual knowledge and the ability to read social cues.¹⁶ There is also debate about whether some human knowledge is tacit and cannot be taught.⁴² Claims that AI will be able to display autonomy have been questioned on grounds that this is a property essential to being human and by definition cannot be held by a machine.¹⁷

ETHICAL AND SOCIAL ISSUES

Many ethical and social issues raised by AI overlap with those raised by data use; automation; the reliance on technologies more broadly; and issues that arise with the use of assistive technologies and ‘telehealth’.

The performance of symptom checker apps using AI, has been questioned. For example, it has been found that recommendations from apps might be overly cautious, potentially increasing demand for unnecessary tests and treatments.¹⁶

RELIABILITY AND SAFETY

Reliability and safety are key issues where AI is used to control equipment, deliver treatment, or make decisions in healthcare. AI could make errors and, if an error is difficult to detect or has knock-on effects, this could have serious implications.⁴³ For example, in a 2015 clinical trial, an AI app was used to predict which patients were likely to develop complications following pneumonia, and therefore should be hospitalised. This app erroneously instructed doctors to send home patients with asthma due to its inability to take contextual information into account.⁴⁴

TRANSPARENCY AND ACCOUNTABILITY

It can be difficult or impossible to determine the underlying logic that generates the outputs produced by AI.⁴⁵ Some AI is proprietary and deliberately kept secret, but some are simply too complex for a human to understand.⁴⁶ Machine-learning technologies can be particularly opaque because of the way they continuously tweak their own parameters and rules as they learn.⁴⁷ This creates problems for validating the outputs of AI systems, and identifying errors or biases in the data.

The new EU General Data Protection Regulation (GDPR) states that data subjects have the right not to be subject to a decision based solely on automated processing that produces legal or similarly significant effects. It further states that information provided to individuals when data about them are used should include “the existence of automated decision-making, (...) meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject”.⁴⁸ However, the scope and content of these restrictions - for example, whether and how AI can be intelligible - and how they will apply in the UK, remain uncertain and contested.⁴⁹ Related questions include who is accountable for decisions made by AI and how anyone harmed by the use of AI can seek redress.³

DATA BIAS, FAIRNESS, AND EQUITY

Although AI applications have the potential to reduce human bias and error, they can also reflect and reinforce biases in the data used to train them.⁵⁰ Concerns have been raised about the potential of AI to lead to discrimination in ways that may be hidden or which may not align with legally protected characteristics, such as gender, ethnicity, disability, and age.⁵¹ The House of Lords Select Committee on AI has cautioned that datasets used to train AI systems are often poorly representative of the wider population and, as a result, could make unfair decisions that reflect wider prejudices in society. The Committee also found that biases can be embedded in the algorithms themselves, reflecting the beliefs and prejudices of AI developers.⁵² Several commentators have called for increased diversity among developers to help address this issue.⁵³

The benefits of AI in healthcare might not be evenly distributed. AI might work less well where data are scarce or more difficult to collect or render digitally.⁵⁴ This could affect people with rare medical conditions, or others who are underrepresented in clinical trials and research data, such as Black, Asian, and minority ethnic populations.⁵⁵

TRUST

The collaboration between DeepMind and the Royal Free Hospital in London led to public debate about commercial companies being given access to patient data.⁵⁶ Commentators have

warned that there could be a public backlash against AI if people feel unable to trust that the technologies are being developed in the public interest.⁵⁷

At a practical level, both patients and healthcare professionals will need to be able to trust AI systems if they are to be implemented successfully in healthcare.⁵⁸ Clinical trials of IBM’s Watson Oncology, a tool used in cancer diagnosis, was reportedly halted in some clinics as doctors outside the US did not have confidence in its recommendations, and felt that the model reflected an American-specific approach to cancer treatment.⁵⁹

EFFECTS ON PATIENTS

AI health apps have the potential to empower people to evaluate their own symptoms and care for themselves when possible. AI systems that aim to support people with chronic health conditions or disabilities could increase people’s sense of dignity, independence, and quality of life; and enable people who may otherwise have been admitted to care institutions to stay at home for longer.⁶⁰ However, concerns have been raised about a loss of human contact and increased social isolation if AI technologies are used to replace staff or family time with patients.⁶¹

AI systems could have a negative impact on individual autonomy: for example, if they restrict choices based on calculations about risk or what is in the best interests of the user.⁶² If AI systems are used to make a diagnosis or devise a treatment plan, but the healthcare professional is unable to explain how these were arrived at, this could be seen as restricting the patient’s right to make free, informed decisions about their health.⁶³ Applications that aim to imitate a human companion or carer raise the possibility that the user will be unable to judge whether they are communicating with a real person or with technology. This could be experienced as a form of deception or fraud.⁶⁴

EFFECTS ON HEALTHCARE PROFESSIONALS

Healthcare professionals may feel that their autonomy and authority is threatened if their expertise is challenged by AI.⁶⁵ The ethical obligations of healthcare professionals towards individual patients might be affected by the use of AI decision support systems, given these might

be guided by other priorities or interests, such as cost efficiency or wider public health concerns.⁶⁶

As with many new technologies, the introduction of AI is likely to mean the skills and expertise required of healthcare professionals will change. In some areas, AI could enable automation of tasks that have previously been carried out by humans.² This could free up health professionals to spend more time engaging directly with patients. However, there are concerns that the introduction of AI systems might be used to justify the employment of less skilled staff.⁶⁷ This could be problematic if the technology fails and staff are not able to recognise errors or carry out necessary tasks without computer guidance. A related concern is that AI could make healthcare professionals complacent, and less likely to check results and challenge errors.⁶⁸

DATA PRIVACY AND SECURITY

AI applications in healthcare make use of data that many would consider to be sensitive and private. These are subject to legal controls.⁶⁹ However, other kinds of data that are not obviously about health status, such as social media activity and internet search history, could be used to reveal information about the health status of the user and those around them. The

Nuffield Council on Bioethics has suggested that initiatives using data that raise privacy concerns should go beyond compliance with the law to take account of people's expectations about how their data will be used.⁷⁰

AI could be used to detect cyber-attacks and protect healthcare computer systems. However, there is the potential for AI systems to be hacked to gain access to sensitive data, or spammed with fake or biased data in ways that might not easily be detectable.⁷¹

MALICIOUS USE OF AI

While AI has the potential to be used for good, it could also be used for malicious purposes. For example, there are fears that AI could be used for covert surveillance or screening. AI technologies that analyse motor behaviour, (such as the way someone types on a keyboard), and mobility patterns detected by tracking smartphones, could reveal information about a person's health without their knowledge.⁷² AI could be used to carry out cyber-attacks at a lower financial cost and on a greater scale.⁷³ This has led to calls for governments, researchers, and engineers to reflect on the dual use nature of AI and prepare for possible malicious uses of AI technologies.⁷³

CHALLENGES FOR GOVERNANCE

AI has applications in fields that are subject to regulation, such as data protection, research, and healthcare. However, AI is developing in a fast-moving and entrepreneurial manner that might challenge these established frameworks. A key question is whether AI should be regulated as a distinct area, or whether different areas of regulation should be reviewed with the possible impact of AI in mind.⁵

Further challenges include the need to ensure that the way AI is developed and used is transparent, accountable, and compatible with public interest, and balanced with the desire to drive UK innovation.⁷⁴ Many have raised the need for researchers, healthcare professionals, and policy-makers to be equipped with the relevant skills and knowledge to evaluate and make the best use of AI.²

THE FUTURE OF AI

In the future, it is likely that AI systems will become more advanced and attain the ability to carry out a wider range of tasks without human control or input. If this comes about, some have suggested that AI systems will need to learn to 'be ethical' and to make ethical decisions.⁷⁵ This is the subject of much philosophical debate,

raising questions about whether and how ethical values or principles can ever be coded or learnt by a machine; who, if anyone, should decide on these values; and whether duties that apply to humans can or should apply to machines, or whether new ethical principles might be needed.⁷⁵

CONCLUSIONS

AI technologies are being used or trialled for a range of purposes in the field of healthcare and research, including detection of disease, management of chronic conditions, delivery of health services, and drug discovery. AI technologies have the potential to help address important health challenges, but might be limited by the quality of available health data, and by the inability of AI to possess some human characteristics, such as compassion.

The use of AI raises a number of ethical and social issues, many of which overlap with issues raised by the use of data and healthcare technologies more broadly. A key challenge for future governance of AI technologies will be ensuring that AI is developed and used in a way that is transparent and compatible with the public interest, whilst stimulating and driving innovation in the sector.

REFERENCES

- 1 See, for example, Engineering and Physical Sciences Research Council, [Artificial intelligence technologies](#).
- 2 US National Science and Technology Council (2016) [Preparing for the future of artificial intelligence](#).
- 3 Royal Society (2017) [Machine learning: the power and promise of computers that learn by example](#).
- 4 The development of AI systems usually involves 'training' them with data. For an overview of different training models, see Nesta (2015) [Machines that learn in the wild: machine learning capabilities, limitations and implications](#).
- 5 House of Lords Select Committee on Artificial Intelligence (2018) [AI in the UK: ready, willing and able?](#).
- 6 Department for Business, Energy & Industrial Strategy (2017) [Policy paper: industrial strategy: the grand challenges](#); Gov.uk (26 April 2018) [Tech sector backs AI industry with multi-million investment](#).
- 7 CBInsights (2017) [AI, healthcare & the future of drug pricing: investment activity, market breakdown, AI in clinical trials](#).
- 8 Gov.uk (22 November 2017) [Autumn budget 2017: 25 things you need to know](#).
- 9 Nuffield Foundation (28 March 2018) [The Nuffield Foundation announces new £5 million Ada Lovelace Institute](#).
- 10 See: <https://www.partnershiponai.org/>.
- 11 IEEE (2018) [The IEEE Global Initiative on ethics of autonomous and intelligent systems](#).
- 12 UNICRI (2017) [UNICRI centre for artificial intelligence and robotics](#).
- 13 Future of Life Institute (2017) [Asilomar AI principles](#).
- 14 Reform (2018) [Thinking on its own: AI in the NHS](#).
- 15 Future Advocacy (2018) [Ethical, social, and political challenges of artificial intelligence in health](#).
- 16 Nesta (2018) [Confronting Dr Robot: creating a people-powered future for AI in health](#).
- 17 European Group on Ethics in Science and New Technologies (2018) [Artificial intelligence, robotics, and 'autonomous' systems](#).
- 18 Science and Technology Committee (Commons) (2018) [Algorithms in decision-making inquiry](#).
- 19 Harrow Council (2016) [IBM and harrow council to bring watson care manager to individuals in the UK](#).
- 20 Alder Hey Children's NHS Foundation Trust (2017) [Welcome to Alder Hey – the UK's first cognitive hospital](#).
- 21 See, for example, Leung MKK, et al. (2016) Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets *Proc IEEE* **104**: 176-97; Science Magazine (7 July 2017) [The AI revolution in science](#).
- 22 O'Mara-Eves A, et al. (2015) Using text mining for study identification in systematic reviews: a systematic review of current approaches *Syst Rev* **4**: 5.
- 23 The Conversation (11 November 2013) [Artificial intelligence uses biggest disease database to fight cancer](#).
- 24 Williams K, et al. (2015) Cheaper faster drug development validated by the repositioning of drugs against neglected tropical diseases *J R Soc Interface* **12**: 20141289.
- 25 Alder Hey Children's NHS Foundation Trust (2016) [Alder Hey children's hospital set to become UK's first 'cognitive' hospital](#).
- 26 Dilsizian SE and Siegel EL (2013) Artificial intelligence in medicine and cardiac imaging *Curr Cardiol Rep* **16**: 441; Future Advocacy (2018) [Ethical, social, and political challenges of artificial intelligence in health](#).
- 27 [Written evidence from Royal College of Radiologists \(AIC0146\) to the House of Lords Select Committee on Artificial Intelligence](#); Hainc N, et al. (2017) The bright, artificial intelligence-augmented future of neuroimaging reading *Front Neurol* **8**: 489.
- 28 Wang D, et al. (2016) Deep learning for identifying metastatic breast cancer *arXiv preprint arXiv:160605718*; Esteva A, et al. (2017) Dermatologist-level classification of skin cancer with deep neural networks *Nature* **542**: 115; Rajpurkar P, et al. (2017) CheXNet: Radiologist-level pneumonia detection on chest x-rays with deep learning *arXiv preprint arXiv:171105225*; Moorfields Eye Hospital (2018) [DeepMind Health Q&A](#).
- 29 See <http://www.ultronics.com/technology/>.
- 30 Bedi G, et al. (2015) Automated analysis of free speech predicts psychosis onset in high-risk youths, *NPJ Schizophrenia*, **1**: 15030; IBM Research (5 January 2017) [IBM 5 in 5: with AI, our words will be a window into our mental health](#).
- 31 Kassahun Y, et al. (2016) Surgical robotics beyond enhanced dexterity instrumentation *Int J Comp Ass Rad* **11**: 553-68.
- 32 Medical News Bulletin (20 January 2017) [Artificial intelligence app Ada: your personal health companion](#); see also: <https://ada.com/>.
- 33 See: <https://www.gpathand.nhs.uk/our-nhs-service>.
- 34 IBM press release (13 Mar 2017) [Arthritis research UK introduces IBM watson-powered 'virtual assistant' to provide information and advice to people with arthritis](#).
- 35 See, for example, <https://chiron.org.uk/>; University of Portsmouth press release (11 September 2014) [Meet Rita, she'll be taking care of you](#). However, some question the suggestion that assistive technologies could relieve pressures on women in particular, see: Parks JA (2010) Lifting the burden of women's care work: should robots replace the "human touch"? *Hypatia* **25**: 100-20.
- 36 Shafner L, et al. (2017) [Evaluating the use of an artificial intelligence \(AI\) platform on mobile devices to measure and support tuberculosis medication adherence](#).
- 37 Moore SF, et al. (2018) Harnessing the power of intelligent machines to enhance primary care *Bri J Gen Pract* **68**: 6-7.
- 38 Jacobsmeyer B (2012) Focus: tracking down an epidemic's source *Physics* **5**: 89; Doshi R, et al. (2017) Tuberculosis control, and the where and why of artificial intelligence *ERJ Open Res* **3**.

- 39 Pirmohamed M, et al. (2004) Adverse drug reactions as cause of admission to hospital *BMJ* **329**: 15-9; Jamal S, et al. (2017) Predicting neurological adverse drug reactions based on biological, chemical and phenotypic properties of drugs using machine learning models *Sci Rep* **7**:872.
- 40 Nuffield Council on Bioethics (2015) [The collection, linking and use of data in biomedical research and health care: ethical issues](#); PHG Foundation (2015) [Data sharing to support UK clinical genetics & genomics services](#); and Reform (2018) [Thinking on its own: AI in the NHS](#).
- 41 Parks JA (2010) Lifting the burden of women's care work: should robots replace the "human touch"? *Hypathia* **25**: 100-20.
- 42 See, for example, Autor D (2014) *Polanyi's paradox and the shape of employment growth*, Volume 20485: National Bureau of Economic Research; Carr, N. (2015) *The glass cage: where automation is taking us*; and Susskind R and Susskind D (2015) *The future of the professions: how technology will transform the work of human experts*.
- 43 Wachter R (2015) *The digital doctor: hope, hype and harm at the dawn of medicine's computer age*. An example of risks posed by hard to detect software error in healthcare is the Therac 25 scandal in 1985-7 when faulty computerised radiation equipment led to accidental overdoses causing six deaths in Canada and the US, see <https://web.stanford.edu/class/cs240/old/sp2014/readings/therac-25.pdf>.
- 44 Caruana R, et al. (2015) Intelligible models for healthcare, in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp1721-30.
- 45 MIT Technology Review (11 April 2017) [The dark secret at the heart of AI](#).
- 46 Burrell J (2016) How the machine 'thinks' *Big Data Soc* **3**:1.
- 47 *ibid*.
- 48 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC ([General Data Protection Regulation](#)).
- 49 PHG Foundation blog (19 September 2017) [The Data Protection Bill: contortionist-like drafting](#); Wachter S, et al. (2017) Why a right to explanation of automated decision-making does not exist in the general data protection regulation *IDPL* **7**: 76-99.
- 50 Dilisizian SE and Siegel EL (2013) Artificial intelligence in medicine and cardiac imaging *Curr Cardiol Rep* **16**: 441; Bird S, et al. (2016) [Exploring or exploiting? Social and ethical implications of autonomous experimentation in AI](#).
- 51 Bird S, et al. (2016) [Exploring or exploiting? Social and ethical Implications of autonomous experimentation in AI](#).
- 52 House of Lords Select Committee on Artificial Intelligence (2018) [AI in the UK: ready, willing and able?](#); Reform (2018) [Thinking on its own: AI in the NHS](#).
- 53 See, for example, The New York Times (25 June 2016) [Artificial intelligence's white guy problem](#); MIT Technology Review (14 February 2018) ["We're in a diversity crisis"](#).
- 54 British Academy and Royal Society (2017) [Data management and use: governance in the 21st century](#).
- 55 See, for example, The Society for Women's Health Research, The US FDA Office of Women's Health (2011) [Dialogues on diversifying clinical trials](#); Popejoy AB and Fullerton SM (2016) Genomics is failing on diversity *Nature* **538**:161; Quartz (10 July 2017) [If you're not a white male, artificial intelligence's use in healthcare could be dangerous](#).
- 56 The Guardian (16 May 2016) [Google DeepMind 1.6m patient record deal 'inappropriate'](#).
- 57 The Guardian (3 November 2017) [Why we can't leave AI in the hands of big tech](#).
- 58 The Conversation (9 January 2018) [People don't trust AI – here's how we can change that](#).
- 59 STAT (5 September 2017) [IBM pitched its Watson supercomputer as a revolution in cancer care. It's nowhere close](#).
- 60 Sharkey A and Sharkey N (2012) Granny and the robots: ethical issues in robot care for the elderly *Ethics Inf Technol* **14**: 27-40.
- 61 *ibid*.
- 62 *ibid*.
- 63 Mittelstadt B (2017) The doctor will not see you now in *Otto P and Gräf E (2017) 3TH1CS: a reinvention of ethics in the digital age?* Otto P and Gräf E (Editors).
- 64 Wallach W and Allen C (2008) *Moral machines: teaching robots right from wrong*.
- 65 Hamid S (2016) The opportunities and risks of artificial intelligence in medicine and healthcare *CUSPE Communications*, summer 2016.
- 66 Cohen IG AR, et al. (2014) The legal and ethical concerns that arise from using complex predictive analytics in health care *Health Aff* **33**: 1139-47.
- 67 Wachter R (2015) *The digital doctor: hope, hype and harm at the dawn of medicine's computer age*.
- 68 Carr N. (2015) *The glass cage: where automation is taking us*; Towards Data Science (24 June 2017) [The dangers of AI in health care: risk homeostasis and automation bias](#).
- 69 Nuffield Council on Bioethics (2015) [The collection, linking and use of data in biomedical research and healthcare: ethical issues](#); Information Commissioner's Office (2018) [Guide to the General Data Protection Regulation \(GDPR\)](#).
- 70 See <https://www.scnsoft.com/case-studies/ibm-qradar-siem-for-a-hospital-with-2000-staff>; IBM Security (2015) [Securing the healthcare enterprise](#).
- 71 Brundage M, et al. (2018) The malicious use of artificial intelligence *arXiv preprint arXiv: 180207228*; Finlayson SG, et al. (2018) Adversarial attacks against medical deep learning systems *arXiv preprint arXiv: 180405296*.
- 72 Yuste R, et al. (2017) Four ethical priorities for neurotechnologies and AI *Nature* **551**:159.
- 73 Brundage M, et al. (2018) The malicious use of artificial intelligence *arXiv preprint arXiv: 180207228*.
- 74 Powles J and Hodson H (2017) Google DeepMind and healthcare in an age of algorithms *Health Technol* **7**: 351-67.
- 75 See, for example, Bostrom N and Yudowsky E (2014) The ethics of artificial intelligence in *The Cambridge handbook of artificial intelligence* Frankish K and Ramsey WM (editors).

Acknowledgments: Thank you to Natalie Banner (Wellcome); Alison Hall, Johan Ordish and Sobia Raza (PHG Foundation); Brent Mittelstadt (Research Fellow at the Oxford Internet Institute, Turing Fellow at the Alan Turing Institute); Ben Moody (techUK); and Reema Patel (Nuffield Foundation), for reviewing a draft of this briefing note.

Published by Nuffield Council on Bioethics, 28 Bedford Square, London WC1B 3JS

May 2018

© Nuffield Council on Bioethics 2018

 bioethics@nuffieldbioethics.org

 @Nuffbioethics

 NuffieldBioethics

www.nuffieldbioethics.org