NHS and Healthcare Data
Debate on 6 September 2018

Summary

This Lords Library Briefing has been prepared in advance of the debate due to take place on 6 September 2018 in the House of Lords on the motion moved by Lord Freyberg (Crossbench), "that this House takes note of NHS and healthcare data and how that data could be used to improve the health of the nation".

This Briefing begins by describing what healthcare data is and the uses to which it might be put. One is to improve decision-making in patient care. Other examples include monitoring both the success of treatments and the performance of care providers. Outside the NHS, data is sought by, for example, academics for research and commercial organisations for monitoring and developing drugs. The Medical Research Council has stated that the UK could be a global leader in the field of health data research.

However, there are also challenges in using healthcare data. Some of these are technological, but there is also widespread concern about privacy and public trust, partly in light of recent breaches of data security. The National Data Guardian for Health and Care oversees these issues.

The use of data in healthcare in England is guided and delivered by a range of organisations. NHS England and the Department of Health and Social Care lead on overall strategy and funding, while the National Information Board guides strategy specifically on data and technology. The main delivery bodies are NHS Digital and Public Health England. However, the roles of these bodies have sometimes overlapped, and some past strategies for data usage are no longer considered appropriate. Steps currently being taken to address these issues include a reorganisation of the roles of NHS Digital and Public Health England. The recently appointed Secretary of State for Health and Social Care, Matt Hancock, has also stated that technology was one of his three "early priorities" and committed £475 million to enhance technology-driven health and care.

Two emerging data-driven technologies are artificial intelligence and genomics. These highlight the potential commercial value of NHS data and there are debates about how this might be realised. Data and technology might also be used to reduce costs in the NHS, with one estimate putting potential savings at 7 to 11.5 percent of healthcare expenditure. Finally, this Briefing provides links to articles describing some specific examples of the uses of data in healthcare.

Chris Smith
31 August 2018

Table of Contents

_____

1. Introduction

The scope for healthcare data to be used to improve health outcomes has been much discussed. It is summarised by the Parliamentary Office of Science and Technology (POST) as follows:

> The NHS holds millions of electronic medical records on the health of the population from birth to death. Increased integration and analysis of these alongside other datasets may provide insights that can improve the understanding and management of the population's health.[1]

However, concerns over its potential use have also been raised, such as in a joint report from Involve, Carnegie UK Trust and Understanding Patient Data, which stated:

> The proliferation of vast amounts of data about who we are and how we live our lives brings significant risk, to our privacy, our autonomy and our sense of identity. Addressing these challenges and identifying where the acceptable trade-offs lie, appears certain to be a complex, high profile public policy priority for many years to come.[2]

This Briefing describes what healthcare data is, how it might be used and some of the challenges in doing so. It summarises the main organisations who operate in and oversee the sector, and describes the current policy position on healthcare data. It then discusses the debate around the use of data-driven emerging technologies in the NHS, including artificial intelligence (AI), 'big data' and genomics. Finally, it provides links to some reports of examples of data-driven healthcare initiatives.

This Briefing discusses health data in the context of the NHS in England, unless otherwise stated. Further information for the other regions of the UK can be found at:

- NHS Research Scotland, 'Scotland's Health Informatics Capabilities', accessed 29 August 2018
- NHS Wales, 'Informatics Service', accessed 29 August 2018
- Health and Social Care Board, eHealth and Care Strategy for Northern Ireland, 3 March 2016

---

[1] Parliamentary Office of Science and Technology, Big Data and Public Health, July 2014, p 1.
[2] Kaela Scott et al, Data for Public Benefit: Balancing the Risks and Benefits of Data Sharing, Involve, Carnegie UK Trust and Understanding Patient Data, April 2018, p 2.

2. Healthcare Data

2.1 What is Healthcare Data?

In July 2014, POST produced a note, *Big Data and Public Health*, on the use of patient health records and other large scale medical and administrative datasets.[3] It set out how some NHS data covers a large proportion of the population, including records from GPs and Hospital Episode Statistics, which cover every attendance, diagnosis and treatment in NHS hospitals in England. It also noted there were other types of data available which covered only subsets of the population, such as genetic information and biobank samples, clinical trials data and public survey results.

Looking further afield, the POST note described how "administrative datasets held by the public sector", such as on school attainment, deprivation, tax and benefit records and births and deaths, could be linked to individuals' medical records. It claimed this could enable "researchers to study the links between health patterns and factors such as education, environment or socio-economic status".[4] It also suggested that there was the potential to link health with non-health sources of data, such as "patient experiences shared on social media platforms and data from supermarkets on consumer habits".[5]

2.2 Uses of Healthcare Data

POST identified three different types of usage of medical records for health improvement.[6] These were:

- **Primary use within the NHS**: usage by medical professionals to make decisions about care.
- **Secondary use within the NHS**: to inform, for example, "commissioning, clinical audit, treatment outcome monitoring, calculation of treatment costs and payment to practices".
- **Secondary use outside the NHS**: data may be used for "local or national public health activities; academic research; and data use by commercial organisations". Commercial organisations might use data for monitoring drug safety and efficacy, and for developing new treatments. POST noted that research often crossed sectoral boundaries, potentially involving the public and private sectors as well as charities and universities.

---

[3] Parliamentary Office of Science and Technology, *Big Data and Public Health*, July 2014.
[4] ibid, p 2.
[5] ibid, p 1.
[6] ibid, pp 2–3.

Looking specifically at patient outcomes, the economics consultancy Volterra distinguished two categories of uses for data analytics in healthcare:

- **Risk stratification and prevention**: data analytics can identify the combination of factors that puts the patient at high risk of developing a chronic condition, allowing for intervention to prevent them from getting ill.
- **Improved treatment outcomes**: personalised medicine can improve early diagnosis and improve quality of care; treatments and outcomes can be analysed in conjunction with patient details in order to maximise the benefit of any treatment.[7]

As to the extent to which healthcare data is being used, in 2017 the Medical Research Council (MRC) brought together information on all organisations working in the field of healthcare data research in the UK.[8] It suggested that recent "considerable" investment across 26 research centres had resulted in "a significant step-change in the UK health data landscape across the spectrum, from molecules to man".[9] The report concluded that:

> The UK has a very real opportunity to be a global leader in health data research and by harnessing this potential could yield huge benefits for our nation's health and wealth.[10]

A response to a parliamentary question in October 2017 stated that there were no limitations on who may apply to use healthcare data and described how requests for such data are handled by the relevant body, NHS Digital.[11] It also set out progress in linking some healthcare datasets as at that date.

2.3 Challenges in Using Healthcare Data

In *Big Data and Public Health*, POST identified two key areas which presented challenges for the use of healthcare data: those around the adequacy of technology and those related to public trust and the legal framework for data access.

Technological Challenges

Regarding technology, POST noted that there was no central database for medical records. Further, because data relating to any individual may come from multiple care settings, a key challenge is to extract the data from each

---

[7] Nick Bosanquet and Ellie Evans, *Sustaining Universal Healthcare in the UK: Making Better Use of Information*, Volterra, September 2014, p 7.
[8] Ekaterini Blaveri, *Mapping the Landscape of UK Health Data Research and Innovation: A Snapshot of Activity in 2017*, Medical Research Council, October 2017.
[9] ibid, p 5.
[10] ibid.
[11] House of Lords, 'Written Question: Medicine: Research', 31 October 2017, HL2194.

location and then link it together.[12] Hence, a common term in recent discussions of healthcare data usage is "interoperability", which has been defined as "the ability of health information systems to work together within and across organizational boundaries in order to advance the effective delivery of healthcare for individuals and communities".[13]

POST also described potential issues around data quality, completeness and accuracy. It gave the example of potentially incorrect codes being used to record illnesses and treatments, or missing or invalid identifiers such as NHS numbers.

Public Trust and the Legal Framework Around Data Access

Regarding public trust and legal usage, POST stated:

> Public trust in the governance of data use is considered key to the continued and expanded use of medical records for health research and management. Surveys suggest that the public is broadly supportive of the use of data for medical research. However, numerous concerns remain. These include: use without consent; use of identifiable data; data security; lack of transparency; potential discrimination by employers or insurers; and access by commercial organisations. Such concerns may result in patients withholding information from healthcare providers, which may be detrimental to the patient and reduce data quality.[14]

It noted that the use of identifiable medical records was "strictly regulated".

The legislative background for the use of medical records was explored as part of a review by Professor Keith McNeil in 2017 (the 'McNeil Review') of Public Health England's handling of healthcare data.[15] The review stated:

> The legal framework governing the use of personal confidential data in health care is complex. It includes the Data Protection Act 1998, the Human Rights Act 1998, NHS Act 2006, the Health and Social Care Act 2012, and the Care Act 2014, as well as the common law duty of confidentiality, the Caldicott principles[16] and a range of national guidance for health and care organisations on data protection and information governance.
>
> The law allows personal data to be shared between those offering treatment and care directly to patients but it protects patients'

[12] Parliamentary Office of Science and Technology, *Big Data and Public Health*, July 2014, p 3.
[13] Health Information and Management Systems Society, 'What is Interoperability?', accessed 16 August 2018.
[14] Parliamentary Office of Science and Technology, *Big Data and Public Health*, July 2014, p 4.
[15] Professor Keith McNeil, *Review of Public Health England's Data Collection and Data Management Functions*, 17 November 2017, p 9.
[16] NHS Digital, 'What are the Caldicott Principles?', accessed 29 August 2018.

confidentiality when data about them are used for other purposes. These "secondary uses" of data are essential if we are to run a safe, efficient, and equitable health service, and achieve the "triple aim" of healthcare provision—better patient outcomes, better patient experience and affordability of healthcare services. For secondary uses where it is not essential to be able to identify patients, data that has been aggregated or anonymised must be used.

The British Medical Association (BMA), the trade union and professional body for doctors and medical students in the UK, has supported the use of data for "secondary" purposes; for example, to assist "administration, planning, audit, commissioning and payment by results".[17] However, it has recognised the challenges in doing so, including keeping patients informed, streamlining the system for handling requests and ensuring that doctors are clear on their legal responsibilities.[18]

The safeguards around the secondary use of patient data, where not de-identified,[19] were strengthened by the introduction of the General Data Protection Regulation (GDPR) on 25 May 2018.[20] On the same day, the national data opt-out for confidential patient information, recommended in the National Data Guardian's 2016 report (discussed further in section 3.7 of this Briefing), was introduced.[21]

Also in accordance with the GDPR, NHS England published a privacy policy in May 2018 (since updated).[22] It included detailed information on why individuals' data is collected and how it is used, and individuals' rights in relation to their data. It states, for example, that confidential patient information can be used by hospital and university researchers, medical royal colleges and pharmaceutical companies researching new treatments, but not for marketing and insurance purposes.[23]

The healthcare thinktank the King's Fund has analysed the implications of these developments.[24] It also considered the implications of what it described as "several high-profile breaches of data security and confidentiality", for example, the Facebook/Cambridge Analytica episode, and a previous attempt at data sharing in the NHS, 'care.data', which was

[17] British Medical Association, 'Secondary Uses of Patient Information', 26 February 2018.
[18] British Medical Association, 'Secondary Uses of Data', 14 March 2017.
[19] 'De-identified' is defined by the Parliamentary Office of Science and Technology as including anonymised, pseudonymised and aggregated data (Parliamentary Office of Science and Technology, Big Data and Public Health, July 2014, p 1).
[20] Understanding Patient Data, 'An Introduction to the GDPR—and what it Means for Patient Data', 6 March 2018.
[21] NHS Digital, 'National Data Opt-out Programme', accessed 16 August 2018.
[22] NHS England, 'Privacy Notice', accessed 28 August 2018.
[23] NHS England, 'Who Uses Your Data', accessed 28 August 2018.
[24] Harry Evans, 'Using Data in the NHS: The Implications of the Opt-out and GDPR', King's Fund, 25 May 2018.

terminated following public concern and criticisms.[25] The King's Fund concluded:

> National policy has to keep a balance between responding to legitimate public concern about the security and confidentiality of data and enabling data to be shared and used by NHS organisations and third parties. It is also essential that NHS national bodies are transparent with the public about how patient data is used.[26]

The King's Fund suggested that the level of opt-outs would be key to the quality and validity of future research, and that NHS England and NHS Digital should keep this under review. It suggested that, in the meantime, these two organisations should "put in place a long-term plan to promote the benefits of NHS organisations and third parties being able to access and use patient data".

A further problem with the handling of patient data was identified in July 2018. Approximately 150,000 objections from individuals about the uses of their data had not been appropriately registered, and these patients' data were therefore shared in ways that they should not have been. The Parliamentary Under Secretary of State for Mental Health and Inequalities, Jackie Doyle-Price, responded to these concerns, stating that the relevant supplier had apologised, and setting out the steps which have been taken to prevent a repeat of such an incident.[27]

3. Public Bodies Involved in Healthcare Data Policy and Practice

The use and oversight of healthcare data in England is the responsibility of several different organisations, whose roles have been defined and developed by a series of policy papers, reviews and strategy documents. This section summarises the key bodies and their main functions. It begins by explaining the involvement in data issues of the oversight bodies NHS England and the Department of Health and Social Care. It then describes the National Information Board, which is the strategic coordination body for data and technology issues in health and care. It goes on to outline the key delivery bodies, NHS Digital and Public Health England, and recent discussions about reorganising their roles. Finally, it describes the role of the National Data Guardian for Health and Care, and mentions some of the other organisations active in the area.

---

[25] Sarah Knapton, 'Controversial £7.5 million NHS Database Scrapped Quietly on Same Day as Chilcot Report', *Telegraph* (£), 6 July 2016.
[26] Harry Evans, 'Using Data in the NHS: The Implications of the Opt-out and GDPR', King's Fund, 25 May 2018.
[27] House of Commons, 'Written Statement: GP Update', 2 July 2018, HCWS813.

3.1 NHS England

NHS England leads the NHS in England and sets its priorities and direction.[28] In 2014, its overall strategy document the *Five Year Forward View* included a section on exploiting the "information revolution",[29] and stated:

> As national bodies we can do more by measuring what matters, requiring comprehensive transparency of performance data and ensuring this data increasingly informs payment mechanisms and commissioning decisions.[30]

In March 2017, NHS England's update to its *Five Year Forward View* included a chapter on "harnessing technology and innovation".[31] Targets and recommendations included:

> Extending mandatory data collections on hospital-acquired infections to more cases, and publishing and learning from locally comparable data on key infections published by Public Health England.[32]

The update emphasised the importance of interoperability between IT systems, not just for immediate patient care, but also to "support the life sciences research strategy".[33] It also described how general practice can be the source of useful data, and reported that the number of GP practices registered on Clinical Research Practice Datalink (a research service using anonymised primary care records) had increased by 25 percent in 2016/17. It then set a target of recruiting an additional 1,800 to 2,000 practices in 2017/18.[34]

3.2 Department of Health and Social Care

The Department of Health and Social Care provides the vast majority of funds for health and social care provision in England.[35]

Considering funding specifically for healthcare data and technology, in February 2016 the then Secretary of State for Health, Jeremy Hunt, announced a £4 billion investment in technology in NHS England.[36] This was confirmed as being "£4.2 billion secured specifically for technology" in the National Information Board's September 2016 *Annual Report*.[37] Press reports

---

[28] NHS England, 'About Us', accessed 29 August 2018.
[29] NHS England, *Five Year Forward View*, October 2014, pp 31–2.
[30] ibid, p 8.
[31] NHS England, *Next Steps on the NHS Five Year Forward View*, March 2017, pp 64–70.
[32] ibid, p 61.
[33] ibid, p 68.
[34] ibid, p 69.
[35] National Audit Office, 'A Short Guide to the Department of Health and NHS England', September 2017, pp 5 and 8.
[36] BBC, 'The Andrew Marr Show', 7 February 2016, (video, reference at 17:25).
[37] National Information Board, *Annual Report*, September 2016, p 4.

provided some breakdowns on the intended use of these funds: for example, that the largest part of the funding was to support the move to a "paperless NHS", while other areas targeted included digital security and data consent, out-of-hospital care and medicines, and digital services for patients.[38]

More recently, in a speech made shortly after his appointment in July 2018 as Secretary of State for Health and Social Care, Matt Hancock, stated that technology was one of his three "early priorities".[39] He announced funding of more than £475 million, which he said would "help jump start the rollout of innovative technology aimed at improving care for patients and supporting staff to embrace technology-driven health and care". He also described the challenges of introducing technology to the NHS, stating that "the small part is finding or inventing the technology. The big part is embedding a culture of always looking for the best possible technology and embracing it".

3.3 National Information Board

The National Information Board (NIB) is a coordinating body made up of representatives of 29 organisations from across the health and care system, supported by a group which provides a "clinical perspective".[40] The NIB described its responsibilities as:

> Developing the strategic priorities for data and technology in health and care to deliver the maximum benefit for all of us, as citizens and patients, and to make appropriate recommendations for investment and action.[41]

The NIB was established as a result of NHS England's *Five Year Forward View* strategy paper.[42] The paper described how, previously, the NHS's approach to information technology "oscillated between two opposite approaches […] neither of which now makes sense". These two approaches were summarised as being highly centralised and highly localised. The highly centralised approach was described as leading to a "lack of local engagement and lack of sensitivity to local circumstances", while the highly localised approach resulted in "systems that don't talk to each other, and a failure to harness the shared benefits that come from interoperable systems".[43] It suggested a more flexible approach should be adopted: key systems should be designed nationally, while local systems would need to meet standards of "interoperability". The NIB was charged with delivering that approach.

---

[38] Pharmafile, 'UK Health Secretary Sets Out £4.2 Billion "Digital NHS" Vision', 8 February 2016.
[39] Department of Health and Social Care, 'Matt Hancock: My Priorities for the Health and Social Care System', 20 July 2018.
[40] National Information Board, *Annual Report*, September 2016, pp 2 and 4.
[41] National Information Board, *Personalised Health and Care 2020: Using Data and Technology to Transform Outcomes for Patients and Citizens*, 13 November 2014, p 3.
[42] NHS England, *Five Year Forward View*, October 2014.
[43] ibid, p 31.

In 2015, the NIB published a series of roadmaps setting out how digital care should be transformed.[44] These covered both extending the uses of data (eg to enhance treatments and to support care professionals in making best use of data and technology) and building public trust (eg by providing digital access to information and clarifying the regulatory framework). In a September 2016 update, the NIB reported that it was "just [at] the beginning of a lengthy process".[45] One of its next steps was entitled 'data outcomes for research and oversight', and described how the NIB would analyse and make available health and social care data in "new and different ways".[46]

3.4 NHS Digital

NHS Digital was set up as the Health and Social Care Information Centre (HSCIC) on 1 April 2013. This followed a 2012 Department of Health paper, *The Power of Information,* which set out a ten-year strategy for using information and technology in healthcare in England.[47] HSCIC changed its name to NHS Digital in July 2016.[48]

In *The Power of Information*, the role of the new body was described as to act as "the focal point for national data collected across the health and care system in England".[49] NHS Digital described its own role as being "the national information and technology partner for the health and care system".[50] More recently, NHS Digital described itself as "the lead national delivery organisation for the technology commitments set out in the National Information Board's strategy".[51]

In July 2017, NHS Digital published an update on its role in delivering the NIB's *Personalised Health and Care 2020* plan.[52] Echoing the NIB's September 2016 update paper, one of its responsibilities was entitled 'data availability for research for outcomes and oversight', described further as seeking to "improve the quality, availability and integrity of health data so that frontline staff, researchers and decision makers are better informed".[53]

---

[44] National Information Board, 'National Information Board's Workstream Roadmaps', 1 October 2015.
[45] National Information Board, *Annual Report*, September 2016, p 6.
[46] ibid, p 27.
[47] Department of Health, *The Power of Information: Putting All of Us in Control of the Health and Care Information We Need*, 21 May 2012. The HSCIC in fact superseded another body of the same name, which had had much narrower responsibilities (Health and Social Care Information Centre, *HSCIC Annual Report and Accounts 2012/13*, 11 July 2013, p 5).
[48] Department of Health, 'HSCIC Changing its Name to NHS Digital', 20 April 2016.
[49] Department of Health, *The Power of Information: Putting All of Us in Control of the Health and Care Information We Need*, 21 May 2012, p 9.
[50] NHS Digital, 'NHS Digital—Our Role in the Health and Care System', accessed 23 August 2018.
[51] NHS Digital, *Business Plan 2017–2018*, August 2017, p 5.
[52] NHS Digital, *Fit for 2020: Report from the NHS Digital Capability Review*, July 2017.
[53] ibid, p 6.

Under the heading of 'information and analysis', NHS Digital stated that it manages over 100 different datasets, and described them falling under the following three headings:

- **Data services**: making data available to customers on health and care services to enable research and analysis by individuals and organisations (including commissioners, providers, arms-length bodies, researchers, the intermediary market etc).
- **Analytical services**: providing standard and bespoke analysis services to customers that either answer their question or provide them with the tools to answer their question.
- **Statistical services**: developing and publishing national, official and other statistical publications that describe health and healthcare across the country.[54]

It continued by outlining the ways in which it would reform its work. These included implementing new datasets, allowing easier access by users to its data, improving the quality of data and extending the data which it collects.[55] The strategy also considered NHS Digital's commercial capabilities, cyber capabilities, internal organisation and how it works with external "customers, stakeholders and partners".[56]

NHS Digital also operates the NHS's Digital Apps Library. According to the NHS England paper which launched the Library, there are three categories of app:[57]

- **"NHS approved" apps**, with a published evidence base, assessed by a process developed with the National Institute for Health and Care Excellent (NICE), demonstrating that they can help a person manage and improve their health;
- **"NHS connected" apps** which have been tested and approved for connection to NHS systems, allowing the user to download information from NHS systems; and
- **"Health apps"** with directories of other health applications from which the user can choose.

By July 2018, the Library included 70 apps.[58]

---

[54] NHS Digital, *Fit for 2020: Report from the NHS Digital Capability Review*, July 2017, p 15.
[55] ibid, p 16.
[56] ibid, pp 18–19.
[57] NHS England, *Next Steps on the NHS Five Year Forward View*, March 2017, pp 64–5.
[58] NHS Digital, 'NHS Apps Library Reaches 70 Apps in Honour of the NHS Birthday', 6 July 2018.

3.5 Public Health England

Public Health England (PHE) was established in 2013,[59] following a commitment made in a 2010 white paper published by the Coalition Government.[60] A 2017 review of its activities, the McNeil Review (see section 3.6 of this Briefing), stated that PHE was formed from more than 100 organisations and described its original role as to "protect and improve the nation's health and wellbeing overall and, in particular, to address health inequalities".[61] The McNeil Review also described the four "critical functions" in PHE's remit letter, one of which was supporting care services with "high quality data" and "data analyses to improve services and outcomes".[62]

Similarly, on PHE's website, one of its six responsibilities is "researching, collecting and analysing data to improve our understanding of public health challenges, and [to] come up with answers to public health problems". [63]

3.6 Reorganising Public Health England and NHS Digital: The McNeil Review

The National Data Guardian for Health and Care's 2016 review (see section 3.7 of this Briefing) suggested an enhanced role for the HSCIC (now NHS Digital). This included its becoming the custodian for all personal confidential data, from which it could then create "comprehensive de-identified data sets".[64]

Building on this, a separate review was established to consider "current arrangements for data collection and data management in PHE and NHS Digital and to evaluate various options for delivering these functions in future".[65] This review was conducted by the first Chief Clinical Information Officer for NHS England, Professor Keith McNeil, and became known as the McNeil Review. It reported in November 2017. It stated that there was:

> An underlying policy objective of developing a central data collection and management function that would operate on behalf of the whole health and social care system to collect and link identifiable health data from all parts of the system. Those datasets could then be made

---

[59] Public Health England, 'About Us', accessed 6 August 2018.
[60] Department of Health, *Healthy Lives, Healthy People: Our Strategy for Public Health in England*, 30 November 2010, Cm 7985, p 2.
[61] Professor Keith McNeil, *Review of Public Health England's Data Collection and Data Management Functions*, 17 November 2017, p 7.
[62] ibid.
[63] Public Health England, 'About Us', accessed 6 August 2018.
[64] National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-outs*, June 2016, p 8.
[65] Professor Keith McNeil, *Review of Public Health England's Data Collection and Data Management Functions*, 17 November 2017, p 20.

available in anonymised form to bona fide organisations for legitimate purposes.[66]

The review noted several issues with the system as it then was, including the excessive burden and cost of data collections and duplicative or overlapping data collections. Respondents to the review suggested that it was a "unique opportunity alongside the portfolio of NIB programmes to enhance data linkages and overall functionality across the system".[67]

The review concluded that "this enhanced functionality can only be provided by NHS Digital in the context of the current information governance landscape and legislation".[68] It noted PHE's argument that "many of those data collections and data management processes (eg disease registration) require specialist clinical oversight and may not be within the current capability of NHS Digital to deliver".[69] It therefore recommended a staged approach for transferring datasets from PHE to NHS Digital. However, it proposed that certain databases, including those supporting "real-time or near real-time provision of patient or population facing public health services", should remain with PHE.[70]

3.7 National Data Guardian for Health and Care

In 2014, Dame Fiona Caldicott was appointed as the first National Data Guardian for Health and Care.[71] According to its website, the Guardian "advises and challenges the health and care system to help ensure that citizens' confidential information is safeguarded securely and used properly".[72]

In 2016, the Guardian, working alongside the Care Quality Commission, published a review of data security, consent and opt-outs.[73] It stated:

> We have a very significant opportunity now to improve the use of data in people's interests, and ensure transparency for the public about when their data will be used and when they can opt out of such usage.[74]

---

[66] Professor Keith McNeil, *Review of Public Health England's Data Collection and Data Management Functions*, 17 November 2017, p 19.
[67] ibid, p 13.
[68] ibid.
[69] ibid, p 20.
[70] ibid, p 5.
[71] Department of Health, *Consultation on the Roles and Functions of the National Data Guardian for Health and Care*, September 2015, p 1.
[72] National Data Guardian, 'About Us', accessed 28 August 2018.
[73] National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-outs*, June 2016.
[74] ibid, p 2.

The review recommended "new data security standards for every organisation handling health and social care information […] intended to support rather than inhibit data sharing".[75] A key proposal was that "people should be able to opt out from personal confidential data being used beyond their own direct care", despite the advantages which accrue from data sharing.[76] This conclusion was based on the overriding importance of earning public trust. However, the review recommended the opt-out should not apply to data which has been anonymised or "de-identified".[77] Thus, for example, commissioners would still be able to use full datasets to analyse whether particular treatments were effective. At the same time, it recommended "the Government should consider introducing stronger sanctions to protect anonymised data".[78]

3.8 Other Organisations

Other organisations operating in the field of healthcare data policy include the following. Please note that as this is a broad and complex field the list below is not comprehensive.

- **National Advisory Group on Health Information Technology in England**
  This group "advises the Department of Health [and Social Care] and NHS England about making the secondary care system more digital".[79] Its 2016 report (the 'Wachter Review') described how digitisation in secondary care has lagged behind that in primary care.[80] It recommended the appointment of a national chief clinical information officer, to "oversee and coordinate NHS clinical digitisation efforts",[81] and set a target that all trusts should have achieved a "high degree of digital maturity by 2023".[82]

---

[75] National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-outs*, June 2016, p 4.

[76] ibid, p 7.

[77] ibid, p 8. 'De-identified' is defined by the Parliamentary Office of Science and Technology as including anonymised, pseudonymised and aggregated data (Parliamentary Office of Science and Technology, *Big Data and Public Health*, July 2014, p 1).

[78] National Data Guardian for Health and Care, *Review of Data Security, Consent and Opt-outs*, June 2016, p 10.

[79] Department of Health, 'Using Information Technology to Improve the NHS', 7 September 2016.

[80] Robert Wachter, *Making Information Technology Work: Harnessing the Power of Health Information Technology to Improve Care in England*, National Advisory Group on Health Information Technology in England, 7 September 2016, p 3. Primary care has been defined as GP practices and community pharmacy, dental and eye services (NHS England, 'Primary Care Services', accessed 7 August 2018); and secondary care as hospitals and specialist care providers (Healthtalk.org, 'Seeing the GP: Advice and Tips for Young People', accessed 7 August 2018).

[81] Robert Wachter, *Making Information Technology Work: Harnessing the Power of Health Information Technology to Improve Care in England*, National Advisory Group on Health Information Technology in England, 7 September 2016, p 38.

[82] ibid, p 5.

- **School for Public Health Research (SPHR) and
  Behaviour and Health Research Unit (BHRU)**
  These organisations were set up as a result of the 2010 white
  paper referred to in section 3.5 of this Briefing.[83] The SPHR is a
  partnership between eight academic institutions, aiming to "build
  the evidence base for effective public health practice". [84] The
  BHRU aims to contribute evidence to efforts to achieve
  behavioural change in areas such as smoking, food, alcohol and
  inactivity.[85]

- **Health Data Research UK (HDR UK)**
  HDR UK was established in 2017 by the Medical Research
  Council and other funders as a new national institute for health
  data science.[86] HDR's stated aim is "to develop and apply cutting
  edge data science approaches in order to address the most
  pressing health research challenges".[87] In February 2018,
  HDR UK announced its first investment of £30 million in
  22 universities and research centres, to be followed by a further
  £24 million in related activities.[88]

4. Emerging Technologies

There has been much debate in recent years about how developing
technologies can be used in healthcare. The 2017 update to NHS England's
Five Year Forward View stated:

> As medicine transforms over the next 20 years, the cutting-edge
> technology in our universities combined with the NHS presents unique
> conditions to create new industries, companies and jobs based in the
> UK and transform the way innovation is delivered to patients, offering
> an attractive place to research and test new treatments in partnership
> with innovators. Potential areas include artificial intelligence, the
> application of genomics to medicine, the development of a range of
> new diagnostic tools, and therapies for conditions that will enable
> more healthy aging.[89]

---

[83] Department of Health, *Healthy Lives, Healthy People: Our Strategy for Public Health in England*, 30 November 2010, Cm 7985, p 2.
[84] National Institute for Health Research School for Public Health Research, 'Welcome to the National Institute for Health Research School for Public Health Research', accessed 28 August 2018.
[85] Behaviour and Health Research Unit, 'About Us', accessed 28 August 2018.
[86] ibid.
[87] Health Data Research UK, 'About', accessed 16 August 2018.
[88] Health Data Research UK, '£54 Million Funding to Transform Health Through Data Science', 7 February 2018.
[89] NHS England, *Next Steps on the NHS Five Year Forward View*, March 2017, p 70.

The King's Fund has also summarised how a range of digital technologies might transform health and care.[90] These include the possibility of hospital-level diagnosis in patients' homes, 'smart pills' which transmit data after being swallowed, and how smartphones might be used to gather analysable data.

This section considers two of these emerging technologies where healthcare data and data analysis play a major role: artificial intelligence (and the closely linked fields of machine learning and the data driven economy) and genomics. One possible consequence of these technologies is that NHS data could have significant commercial value, and this section also looks at what this might be, how it might be realised and how it could lead to additional security risks. Finally, it considers how technology and use of data might produce cost savings in the NHS.

4.1 Artificial Intelligence, Machine Learning and the Data Driven Economy

In its 2017 *Industrial Strategy*, the Department for Business, Energy and Industrial Strategy (BEIS) defined artificial intelligence (AI) as "technologies with the ability to perform tasks that would otherwise require human intelligence, such as visual perception, speech recognition, and language translation".[91] BEIS linked AI to "machine learning", which it described as "a type of AI that allows computers to learn rapidly from large datasets without being explicitly programmed". It also related both these concepts to the "data driven economy", in which value and insights can be extracted rapidly from large-scale, connected datasets—also known as 'big data'.

The *Industrial Strategy* committed to making the UK a "global centre for artificial intelligence and data-driven innovation".[92] Regarding healthcare specifically, it stated the Government would "invest to enhance the power of health data to diagnose life-changing diseases at the earliest possible stage and develop precision treatments to cure them".

The 2018 report of the House of Lords Artificial Intelligence Committee included a chapter on healthcare and AI.[93] The report stated that most witnesses to the Committee believed AI presented "significant opportunities" for the healthcare sector, although there was at least one dissenting voice who believed that claims made for healthcare AI were "overblown".[94] It provided some examples of existing uses of AI, such as in

---

[90] Cosima Gretton and Matthew Honeyman, 'The Digital Revolution: Eight Technologies that will Change Health and Care', King's Fund, 1 January 2016.

[91] Department for Business, Energy and Industrial Strategy, *Industrial Strategy: Building a Britain Fit for the Future*, November 2017, p 37.

[92] ibid, p 38.

[93] House of Lords Artificial Intelligence Committee, *AI in the UK: Ready, Willing and Able?*, 16 April 2018, HL Paper 100 of session 2017–19, pp 87–94.

[94] ibid, p 87.

interpreting scans and x-ray images, and took evidence on possible future uses, both clinical and administrative.[95]

The Committee also considered whether the NHS was well positioned to take advantage of AI. The Committee concluded that it was "concerned by the NHS's lack of organisational preparedness to embrace new technology".[96] It cited other sources which stated that "there is a worrying absence of a credible strategy to encourage the uptake of innovation and technology at scale across the NHS", and that "the digital revolution has largely bypassed the NHS". In a similar vein, in evidence to the Committee, the chair of the DeepMind Independent Review Panel (see section 4.3 of this Briefing) described the arrangements between private companies and NHS Trusts as "very chaotic".[97]

4.2 Genomics

Genomics is the studying of genomes and the sequencing of DNA.[98] Its potential for improving healthcare outcomes in the UK was summarised by the House of Commons Science and Technology Committee in April 2018:

> Genomics has the potential to transform medical care across the world and in the NHS as well as dramatically improve patient outcomes in the UK and globally. The UK is considered to be a world leader in the development of genomics.[99]

The Committee commended the "government policies and on-the-ground initiatives" which had led to this "internationally enviable position", but suggested that "much more needs to be done to achieve the full potential that genomics promises".[100] In particular, it recommended continued investment in digital infrastructure and ensuring that the NHS workforce was prepared for the impact of genomics. It described public education about both the opportunities presented by genomics, and the data protection rules in place, as "essential".[101] The Committee called for a tightening up of the consent framework (by which patients agree to have their data used) for the Genomic Medicine Service.[102]

---

[95] House of Lords Artificial Intelligence Committee, *Artificial Intelligence in the UK: Ready, Willing and Able?*, 16 April 2018, HL Paper 100 of session 2017–19, pp 87–8. For another description of the existing and potential future uses of AI in medicine, see: *Economist* (£), 'Artificial Intelligence Will Improve Medical Treatments', 7 June 2018.

[96] House of Lords Artificial Intelligence Committee, *AI in the UK: Ready, Willing and Able?*, 16 April 2018, HL Paper 100 of session 2017–19, p 92.

[97] ibid, p 91.

[98] Genomic.org.uk, 'What is Genomics?', accessed 14 August 2018.

[99] House of Commons Science and Technology Committee, *Genomics and Genome Editing in the NHS*, 20 April 2018, HC 349 of session 2017–19, p 3.

[100] ibid.

[101] ibid, pp 3–4.

[102] ibid, p 4.

The Committee also recognised the commercial value of genome information. It called on Genomics England to maximise that value, while continuing to provide industrial and academic access to the data.[103] It also encouraged integration of genomic data with other NHS data.

Finally, the Committee briefly considered genome editing, a technology which it said has "ethical considerations, some of which are the subject of particular debate". It noted an existing strong regulatory regime in the area, and called for close monitoring of the development of the technology.

The 2017 update of NHS England's *Five Year Forward View* described the NHS's "genomics capability".[104] It stated that 13 Genomic Medicine Centres had already been established, and that they had collected "over 25,000 cancer and rare disease samples" and sequenced over 20,000 whole human genomes.[105] It went on to set targets of "collecting 50,000 samples in 2017/18, rising to 90,000 in 2018/19", and it estimated that by the end of the 2018 calendar year, 100,000 whole human genomes will have been sequenced. The update also expressed an aim of developing a "genomic medicine service for England, underpinned by a new consolidated genomic laboratory infrastructure and facilitated by informatics, data and analytical platforms".

According to Genomics England, as at the latest update on 6 August 2018, 75,552 genomes had been sequenced.[106]

4.3 Commercial Value of NHS Data and Data Security

The House of Lords Artificial Intelligence Committee report highlighted the possible commercial value of NHS data, particularly given its near-complete coverage of UK citizens and its long history.[107] It stated that "this data could be of immense value to artificial intelligence researchers".[108] However, a witness to the Committee, Nicola Perrin from the Wellcome Trust, described the challenges with monetising the data. She said the public "do not like the idea of the NHS selling data, but they are even more concerned if companies are making a profit at the expense of both the NHS and patients".[109] The Committee went on to suggest an alternative model whereby AI companies might provide their services for free in return for access to data.

---

[103] House of Commons Science and Technology Committee, *Genomics and Genome Editing in the NHS*, 20 April 2018, HC 349 of session 2017–19, pp 4–5.
[104] NHS England, *Next Steps on the NHS Five Year Forward View*, March 2017, p 69.
[105] ibid.
[106] Genomics England, 'The 100,000 Genomes Project by Numbers', accessed 17 August 2018.
[107] House of Lords Artificial Intelligence Committee, *AI in the UK: Ready, Willing and Able?*, 16 April 2018, HL Paper 100 of session 2017–19, pp 88–90.
[108] ibid, p 88.
[109] ibid, p 89.

One example of such an arrangement is between DeepMind Health, a company within the Google group, and the Royal Free London NHS Foundation Trust.[110] This agreement has attracted some comment and criticism: for example, in a House of Lords debate on the Data Protection Bill in May 2018, Lord Mitchell (Crossbench) said:

> Like knights on white chargers, DeepMind has financed the digitisation of millions of patients' data and produced algorithms that are already making a major contribution to improving difficult-to-diagnose conditions. It has cost the Royal Free next to nothing and, unsurprisingly, its staff are over the moon. What they do not realise is that the algorithms produced by DeepMind have international value and will be monetised all over the world for the benefit of Google, not of our NHS.
>
> DeepMind and companies like it are swarming all over the NHS. For my part, to put it bluntly, I want to stop them gathering the benefits of our data on the cheap.[111]

The Information Commissioner also found that the form of this arrangement was in breach of the Data Protection Act 1998, finding "several shortcomings in how the data was handled", including that "patients would not have reasonably expected their information to have been used in this way".[112]

In 2016, DeepMind had established a "panel of independent reviewers" to oversee its activities. The latest annual report of the panel was published in June 2018.[113] As well as assessing the company's specific workstreams and general progress, the report set out twelve principles according to which DeepMind, or similar companies, should operate.[114]

In contrast to the DeepMind model, Lord Drayson (Labour), who has recently floated a digital healthcare company, Sensyne Health, has argued that the NHS should benefit financially at least to some degree from its data. Sensyne's agreements with hospitals provide them with both "an equity stake in the company and a share of the revenues that arise from any discoveries".[115]

---

[110] DeepMind Health, 'DeepMind Health and the Royal Free', accessed 28 August 2018.
[111] HL Hansard, 14 May 2018, cols 428 and 430.
[112] Information Commissioner's Office, 'Royal Free—Google DeepMind Trial Failed to Comply with Data Protection Law', 3 July 2017.
[113] DeepMind Health, 'DeepMind Health Independent Review Panel Annual Report', June 2018.
[114] ibid, p 7.
[115] Paul Drayson, 'The NHS Needs to Cash In Its Data', Telegraph (£), 19 August 2018.

The fact that NHS data may have considerable commercial value means that, according to a report by thinktank Parliament Street, "it could easily be seen as a goldmine for hackers".[116] The report described a number of instances of data loss, including missing paper documents, unavailable electronic records and cyberattacks. It recommended the abolition of handwritten records in hospitals and the introduction of "patient identity protocols" to ensure that all information is properly captured and stored digitally.[117]

4.4 Potential Cost Savings from Data and Technology

In June 2015, the then National Director for Patients and Information at NHS England, Tim Kelsey, was reported as estimating the possible cost savings from digitising the NHS at between £8.3 billion and £13.7 billion.[118] This compares to total health spending in England in 2017/18 of approximately £125 billion.[119]

The estimated savings were said to be based on an unpublished report by consultants McKinsey and Company for NHS England. However, in 2016, McKinsey did publish a report which estimated that for "one large OECD country" the implementation of existing digital technologies "could reduce its healthcare expenditure between 7 and 11.5 percent".[120] These figures would approximately correspond to those quoted above for England.

The McKinsey report contained a breakdown of the savings, suggesting that of the 7 to 11.5 percent:

- 3 to 6 percent would come from "seamless data exchange" (patient "self-service", the benefits of electronic records in reducing examinations and improving clinical decision making, and the savings from remote monitoring and mobile healthcare provision);
- 1 to 1.5 percent would come from "transparency and advanced analytics" (performance based provider incentives, and better provision of public information on healthy and risky behaviours); and
- 3 to 4 percent would come from "process digitisation and automation" (administration, "patient flow management", e-prescriptions, online booking etc).[121]

---

[116] Parliament Street, *NHS Data Security: Protecting Patient Records*, August 2018, p 2.
[117] ibid, p 8.
[118] Thomas Meek, '*Department of Health Refuses to Release IT Savings Report*', Digital Health News, 13 August 2015.
[119] Full Fact, '*Spending on the NHS in England*', 21 June 2018.
[120] Thomas London and Penelope Dash, '*Health Systems: Improving and Sustaining Quality Through Digital Transformation*', McKinsey and Company, August 2016.
[121] ibid.

A further example of a potential cost saving was described in NHS England's *Five Year Forward View*. It stated that:

> The NHS has the opportunity radically to cut the costs of conducting Randomised Controlled Trials (RCTs), not only by streamlining approval processes but also by harnessing clinical technology. We will support the rollout of the Clinical Practice Research Datalink, and efforts to enable its use to support observational studies and quicker lower cost RCTs embedded within routine general practice and clinical care.[122]

5. Examples of Data-<u>D</u>riven Healthcare Initiatives

The following articles provide some examples of the use of data in healthcare:

- University College London Hospital, 'Revolutionising Healthcare with AI and Data Science: UCLH and The Alan Turing Institute Announces Breakthrough Partnership Today', 21 May 2018
- Alan Turing Institute, 'New Research Shows Machine Learning Could Significantly Augment Clinical Decision-making in Cystic Fibrosis Care', 2 August 2018
- Tom Whipple, 'Machines Faster than Humans at Diagnosing Brain Injuries', *Times* (£), 14 August 2018
- Paul Drayson, 'The NHS Needs to Cash In Its Data', *Telegraph* (£), 19 August 2018

---

[122] NHS England, *Five Year Forward View*, October 2014, p 33.