Manufacturing the future

Could healthcare data help rebalance the UK’s economy?

Eleonora Harwich
Annemarie Naylor

February 2020 #healthandwealth
About

This essay is the fruit of a collaboration between Future Care Capital and Reform. The arguments and any errors that remain are the authors’ and the authors’ alone.

Future Care Capital

Future Care Capital is a charity which undertakes research and engages in practical projects to advance ideas that will help shape future health and social care policy and deliver better outcomes for individuals living in the UK.

Beginning life as the National Nursery Examination Board in 1945, the charity has evolved throughout its 70-year history and we continue to have Her Majesty the Queen as our Royal Patron.

Our work currently takes as its focus the opportunities and challenges that data, technology, innovation, scaling and investment represent in respect of health and care.

Future Care Capital is a registered charity, charity no. 036232.

Reform

Reform is established as the leading Westminster think tank for public service reform. We are dedicated to achieving better and smarter public services. Our mission is to set out ideas that will improve public services for all and deliver value for money.

We work on core sectors such as health and social care, education, home affairs and justice, and work and pensions. Our work also covers issues that cut across these sectors, including public service design and delivery and digital public services.

We are determinedly independent and strictly non-party in our approach.

Reform is a registered charity, the Reform Research Trust, charity no.1103739.

About Reform Perspectives

Reform Perspectives – These are long essays, akin to a newspaper long reads, showcasing a particular opinion or angle on a policy issue. These are meant to be thought provoking and inject new and bold policy thinking into the public sector debate.
Manufacturing the future

The Government’s forthcoming budget represents a critically important opportunity to kickstart the UK’s post-Brexit economy and should focus on arresting the cycle of unequal productivity and wealth disparities between regions and the devolved nations.

The former Minister for Universities, Science, Research and Innovation, the Rt Hon Chris Skidmore MP, recently underlined the Government’s ambition to make the UK a “global science superpower” through increased investment in research and development. He also signalled that the high concentration of such investment in the ‘golden triangle’ which straddles Oxford-Cambridge-London is liable to be disrupted, as the Government seeks to develop more of a place-based approach to research funding. This chimes with the Prime Minister’s commitment to a ‘One Nation’ government and his ‘levelling up’ agenda. It also points towards significant challenges, as well as opportunities, that could flow from the UK’s health and life science industries – a sector worth an estimated £75 billion and which currently employs 250,000 people nationwide.

One of the key catalysts of the fourth industrial revolution is data, and The Life Sciences Industrial Strategy highlights the crucial role that healthcare data is expected to play in growing the sector in years to come. In particular, the Government expects the UK’s unique longitudinal healthcare dataset to play a pivotal role in the discovery of new treatments, better diagnosis and improved efficiency for the healthcare system.

Anticipating the future benefits that could be derived from third party use of data collected by health and care organisations is incredibly difficult. It could lead to gross over- or under-estimation because we do not know what the treatments and technologies of tomorrow will look like, or what the need and demand for them might be. Nonetheless some have speculated that the 55 million patient records the NHS holds could be worth several billion pounds in commercial revenues, efficiency savings and wider economic benefits. For example, Ernst and Young estimate that the realisation of £9.6 billion per annum in benefits could be unlocked from patient data alone, and there is further potential to harness non-patient or ‘operational’ datasets.

Many of the insights and innovations that researchers and clinicians anticipate will be created in partnership with private sector organisations because companies can take on the cost and risk of developing new treatments. They can also attract highly skilled labour to develop sophisticated algorithms to analyse the petabytes of data the healthcare system holds. And yet the UK currently lacks a coherent strategy setting out how any public investment should be targeted to maximise the value that flows from data-driven innovation for patients, the health and care system, higher education institutions and

---

commercial partners. Nor is it clear how Government means for such value to be apportioned or shared among these stakeholders.\textsuperscript{7}

Unsurprisingly, there are growing calls for a concerted effort on the part of Government to harness the value of such data through a national strategy to ensure that all parts of the country receive a fair share of any benefits derived from third party access to healthcare data.\textsuperscript{8}

As his first Budget statement fast approaches, the new Chancellor of the Exchequer is faced with the challenge of financing the Government’s ‘levelling up’ agenda as well as with fulfilling the promise to bring about the “the single biggest increase in public R&D spending by any government in the post-war period”.\textsuperscript{9} The Government must ensure that this wave of investment delivers genuinely transformational change for so-called ‘left behind’ communities while boosting the UK economy overall. Strategic deployment of publicly controlled data should, then, be at the forefront of the Chancellor’s mind.

The price of the status quo

The NHS “holds millions of electronic medical records on the health of the population from birth to death” as well as extensive administrative datasets capturing, for example, the number of operations being carried out in the operating theatres of a hospital.\textsuperscript{10} Yet there is currently no systematic way in which to understand what the demand for this data from third parties looks like, as there is no official ledger about data access requests made of health and care organisations in the UK.

Primary research undertaken by Future Care Capital found that there are geographical disparities in the demand for such data as well as differing interest in datasets controlled by primary, secondary and community health and care Trusts.\textsuperscript{11} It issued a Freedom of Information (FOI) request to 227 NHS trusts and received 192 responses. One in five trusts who responded to the request reported that they had no record of the data access requests they received for research or commercial purposes.\textsuperscript{12} Trusts in the West Midlands, London and the North West received more requests than other regions and, overall, requests mapped to the nearby presence of a Russell Group university;\textsuperscript{13} although this analysis does not capture well the demand experienced by larger Trusts and, in particular, those which form the Shelford Group.\textsuperscript{14} Understanding what the demand for

---

\textsuperscript{7} Eleonora Harwich and Rose Lasko-Skinner, Making NHS Data Work for Everyone (Reform, 2018); Saira Ghafur et al., NHS Data: Maximising Its Impact on the Health and Wealth of the United Kingdom, 2020; Annemarie Naylor and Joselene Magnusson, Taking Next Steps to Harness the Value of Health and Care Data (Future Care Capital, n.d.);
\textsuperscript{8} Annemarie Naylor and Emily Jones, Intelligent Sharing: Unleashing the Potential of Health and Care Data in the UK to Transform Outcomes (Future Care Capital, n.d.); ibid.; Harwich and Lasko-Skinner, Making NHS Data Work for Everyone.
\textsuperscript{9} The Rt Hon Chris Skidmore MP, ‘Levelling up Research and Innovation Right across the United Kingdom’.
\textsuperscript{10} Peter Border, Big Data and Public Health (Parliamentary Office of Science and Technology, 2014), 1.
\textsuperscript{11} Naylor and Magnusson, Taking Next Steps to Harness the Value of Health and Care Data, 13–23.
\textsuperscript{12} Ibid., 13.
\textsuperscript{13} Ibid., 15.
\textsuperscript{14} Shelford Group, ‘Shelford Group Members’, Shelford Group, n.d.; (The Shelford Group includes: Cambridge University Hospitals NHS Foundation Trust, Guy’s and St Thomas’, Imperial College Healthcare, King’s College Hospital, Manchester University, Newcastle upon Tyne Hospitals, Oxford University Hospitals, Sheffield Teaching Hospitals, University College London Hospitals, University Hospitals Birmingham).
different types of healthcare data looks like matters as it is the ‘fuel’ for data-driven innovation.

Related to this, research by Reform has suggested that there is a risk that if all partnerships operate at Trust-level only, the Trusts in question will benefit, leaving other Trusts to miss out on products or services and, potentially, financial benefits. In the long run this could create imbalances in the healthcare system as a whole, contradicting the principle that the NHS is meant to “maximise [its] resources for the benefit of the whole community, and make sure nobody is excluded”.

Diverting the commercial value generated at a local level to a centrally controlled entity for redistribution could negatively impact established incentives to innovate, but retaining value where it is disproportionately concentrated runs the risk of exacerbating inequalities in health and wealth.

International context

The UK’s artificial intelligence (AI) and Life Sciences ecosystems do not operate in ‘splendid isolation’ – independent of global investment flows that continue to turbo-charge the data-driven healthcare revolution. Therefore, in addition to regional considerations, the Government must take heed of what its international counterparts are doing in formulating and incentivising any forward plan of action. If it is successful, this will ensure the UK retains its competitive advantage in key sectors and, crucially, is well-placed to anticipate new and emergent industries that might flow from them.

China has made it a clear priority to harness big data in healthcare in order to tackle its unevenly distributed medical resources. It is building huge regional data centres – aggregating data about its citizens such as medical records, insurance claims and other information. It is also investing significant funds – $9 billion dollars are earmarked for the period 2016-2031 – to sequence and analyse genomes. This will increase the scale of healthcare records (1.4 billion population) as well as the scope of data made available for research and innovation as it seeks to develop the treatments and technologies of the future.

---

19 Mak, ‘Breakingviews - Really Big Data Gives China Medical AI Edge’.
20 Ibid.
Before data can be used in the development of new treatments or to train algorithms to diagnose health conditions, however, it needs to be cleaned, annotated and curated.21 The Economist recently reported that one of the key factors behind China’s AI boom is data labelling companies, such as Megvii, which is reported as having spent $31 million cleaning and curating data over the past three-and-a-half years, because data is an inherently messy business and machines require rigour.22 Megvii’s ‘data factories’ are said to employ upwards of 300,000 data labellers and are often located in its poorest provinces.23 Wage arbitrage allows these companies to pay higher wages than the average wage in those areas, while still paying less than would be the case in more affluent regions and urban conurbations.24

China is not the only place to pursue such a model. In India, labelling companies like iMerit and Samasource are operating in a similar vein. iMerit has an “all-female training facility near Kolkata” which has “become a specialist centre for computer vision labelling”.25 Fifty per cent of the company’s workforce is female and most herald from low-income Indian families.26

Both China and India are taking advantage of the fact that data need not be constrained by geography but, instead, can be accessed remotely or moved around. They also do not have the same high levels of concern about cyber security that the UK has. As such, innovative companies in India and China can more readily locate in the most deprived provinces, where they often also benefit from government subsidies and other incentives – some of which are not dissimilar to those which have underpinned Enterprise Zones in the UK since 2012.27 The legal and regulatory frameworks which currently apply to third party (e.g. private companies or research institutions) healthcare data access and usage in the United Kingdom are markedly different.

‘An AI assembly line’

AI is reliant upon high-quality data which is correctly labelled and helpfully structured28 and, for now, data clean-up, annotation and curation remains relatively labour-intensive.

It is for this reason that Flatiron Health was acquired by Roche for $1.9 billion in 2018 – a value due, in large part, to the unique, high-quality dataset it had invested in producing.29

22 The Economist, ‘China’s Success at AI Has Relied on Good Data | Technology Quarterly’, The Economist, 2 January 2020.
23 Ibid.
24 Ibid.
26 Ibid.
28 House of Lords Select Committee on Artificial Intelligence, AI in the UK: Ready, Willing and Able? Report of Session 2017–19; Professor Dame Wendy Hall and Jérôme Pesenti, Growing the Artificial Intelligence Industry in the UK (Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy, 2017); Eleonora Harwich and Kate Laycock, Thinking on Its Own: AI in the NHS (Reform, 2018).
29 Roche, ‘Roche to Acquire Flatiron Health to Accelerate Industry-Wide Development and Delivery of Breakthrough Medicines for Patients with Cancer’, Press Release, n.d.; Reenita Das, ‘The Flatiron Health Acquisition Is A Shot In The Arm For Roche’s Oncology Real-World Evidence
The company’s approach to data curation was to employ qualified healthcare professionals such as nurses and certified tumour registrars to curate and correctly label medical information.30 The company provides “centralized training” for its data curation workforce and has “a formal quality oversight process”.31 As highlighted by the Vice President of Oncology at Flatiron Health, their approach is to turn “electronic health records into an evidence source can be thought of as a manufacturing process”.32 The company manufactures “research-grade (and ultimately, “regulatory-grade”) de-identified real-world data sets from electronic health records to test hypotheses and generate evidence”.33

At first, both China’s and India’s data labelling industry focused on relatively simple labelling tasks, but they are now moving into the clean-up and curation of much more complex types of data including medical images.34 It is unclear exactly what type of training these medical data labellers receive or how much more they are paid compared to their less skilled counterparts.

The curation of medical data takes skill and knowledge of the context in which it was collected. However, if the process can be likened to manufacturing, are there parts of the medical data curation conveyor belt that could be delegated to people with no prior knowledge of healthcare but who could be progressively upskilled? There are also tasks which are inherently difficult for machines to undertake – for now – like deciphering handwritten notes. Could that become someone’s job?

**Lessons for the UK**

With this in mind, investment in R&D, new technologies and industries could be used as a stimulus to revitalise local economies and regions which continue to lag behind their more productive (and, arguably, over-heated) counterparts in London and the South East.35 Richard Jones, a professor of physics at the University of Sheffield, has argued that research investment is a major tool at Government’s disposal which could be strategically deployed to help rebalance the UK’s economy.36

However, a closer look at the UK’s current life sciences innovation supply chain shows a concentration of ‘excellence’ centres, rather than a place-based approach which would benefit the locations that the Government has pledged to support. The same can also be said of NHS Trusts which are at the forefront of medical research and technological

---

31 Ibid.
32 Ibid.
33 Ibid.
34 Mak, ‘Breakingviews - Really Big Data Gives China Medical AI Edge’; The Economist, ‘China’s Success at AI Has Relyed on Good Data | Technology Quarterly’; Murgia, ‘AI’s New Workforce: The Data-Labelling Industry Spreads Globally’.
35 Morgan, ‘Rebalancing UK Research Funding Regionally: How Could It Happen?’
innovation, although there are notable exceptions such as Liverpool and the work of the Connected Health Cities initiative.\textsuperscript{37}

How then should the Government invest in infrastructure, R&D and healthcare data to ‘level-up’ local economies and increase the health and wealth of the nation?

**Physical infrastructure**

It will take access to land and buildings. The UK Government has already established Enterprise Zones and, more recently, Life Science Opportunity Zones which benefit from dedicated facilities that underpin research, innovation and commercial activities.\textsuperscript{38} To replicate the ‘data factory’ model which is propelling the digital manufacturing boom in China and India in order to harness the potential of healthcare data, the Government would need to consider related investments in Digital Innovation Hubs – which are meant to provide “responsible and safe access to health data, technology and science, research and innovation services”.\textsuperscript{39}

The Government would also need to weigh up the relative merits of outsourcing less specialist data clean-up work to data factories overseas and, in particular, attendant data protection, cybersecurity and environmental sensitivities, were it to favour investment in more specialist annotation and curation services incubated and headquartered in the UK.

Investing in towns like Southport in the North West, which already benefits from proximity to supercomputing power as well as the high-speed fibre optic cabling needed to work with big datasets for AI, could have a significant impact on such local economies.\textsuperscript{40}

Currently, Southport (see Figure 1) has been identified as a ‘left-behind area’ as measured by OCSI’s Community Needs Index – which covers 19 indicators from the presence of civic assets like libraries to measures of connectedness like the strength of the local job market. Could it become home to the country’s healthcare data manufacturing or services industry?


\textsuperscript{38} Office for Life Sciences, ‘Life Sciences Opportunity Zones: How to Apply’, Webpage, GOV.UK, 3 July 2019; Matthew Ward, *Enterprise Zones*.

\textsuperscript{39} Health Data Research UK, ‘Infrastructure’, Webpage, HDRUK, n.d.

\textsuperscript{40} Shane Farmer, ‘Digital Infrastructure Plan Announced for Liverpool City Region’, Webpage, Liverpool City Region Combined Authority, 15 January 2018.
A skilled workforce

It will also mean ensuring the sector has the right skills. Low-skilled and low-wage jobs are not going to cut it if the Government is serious about ‘levelling up’. It needs to look at the industries of the future and consider how it can progressively upskill the workforce in targeted geographies.

In the short-term, this is liable to require a step-change investment in Further Education. It would also take a concerted effort on the part of employment agencies to proactively recruit people to training opportunities and next-generation jobs they might not otherwise have considered.

In the medium to long-term, the UK might opt to evolve and specialise in high-end data curation services through automated processes that are less labour intensive. But, by then, it would benefit from a workforce equipped to support the digital transformation of our healthcare system. This might, in turn, benefit the Department of Health and Social Care’s workforce planning efforts and the work of the new Minister with responsibility for workforce issues, the Rt Hon Helen Whatley MP, in developing the upcoming NHS People Plan.41

In addition, to the recommendations made in the Topol Review – which sought to “prepare the healthcare workforce, through education and training, to deliver the digital future”42 – the Department should consider some of the key tenets outlined in Germany’s upskilling strategy known as ‘Industry 4.0’.43 It is a concerted plan designed in partnership between

---

41 NHS Improvement, Interim NHS People Plan, 2019.
43 Centre for New Economy and Society Insight, Towards a Reskilling Revolution Industry-Led Action for the Future of Work (World Economic
industry, unions and educational institutions which “prepares their employees for emerging skills and workplace changes”.\textsuperscript{44} Taken in conjunction these documents could strengthen the upcoming NHS People Plan.

As highlighted in a recent report from the Institute of Global Health Innovation, “senior healthcare leaders have identified funding as the main barrier to the digital transformation of the NHS and estimate the investment to be in the order of £3 billion”.\textsuperscript{45} If the Chancellor agrees on Budget day to a significant investment in enhancing the quality of healthcare data, undertaking the curation work in the UK would be more costly than outsourcing it to other countries where labour is cheaper and employment standards are lower. Although the UK would not be able to compete in terms of the volume and speed at which data can be curated in China, we could compete on the quality of curation via an upskilled workforce.

The Chancellor will have to weigh these options: on the one hand it will take some time to upskill the country’s workforce, but on the other, there would be considerable wider economic benefits from investing in home-grown talent.

**Capital investment**

Cleaning up healthcare data is, by definition, reliant upon a robust, speedy and secure data infrastructure. Ongoing investment will therefore be needed. The Government is committed to cloud-based solutions and NHSX is busy deploying staff to work alongside commissioners of tech products and services at Trust level in order to obtain value for money. It is also said to be mooting a minimum level of spend on Information Communication Technologies. But this is liable to further squeeze already challenging revenue budget allocations.

Meanwhile, Prof Eerke Boiten, Director of the Cyber Technology Institute at De Montfort University, Leicester, has indicated that the CeLSIUS facility operated by University College London is the most secure set-up for healthcare data access by third parties\textsuperscript{46} – with no cloud-based solution currently offering the same level of security. Could a secure and high-speed link between data factories for (initially) less skilled data clean-up work in ‘left-behind’ communities and ‘excellence’ centres for more specialist healthcare data annotation be established?

Could the Chancellor, for example, incentivise the use of Network Rail’s fibre optic network – already connecting cities from Manchester to York\textsuperscript{47} – to connect the likes of Southport with Oxford? This might need to happen within the organisational confines of the NHS in the interests of cybersecurity, which would herald a real digital expansion.

\textsuperscript{44} Centre for New Economy and Society Insight, Towards a Reskilling Revolution Industry-Led Action for the Future of Work, 24.
\textsuperscript{45} Ghafur et al., NHS Data: Maximising its Impact on the Health and Wealth of the United Kingdom.
\textsuperscript{47} Network Rail, ‘Connecting Telecom Services’, Webpage, Network Rail, n.d.
Alternatively, this could be done by the private sector or by higher education institutions using established research networks like Jisc’s Janet network.48

**Safeguards for individual rights and regulation**

The UK isn’t China, India or the United States, and the public will, rightly, demand that any investments in the healthcare data manufacturing industries of tomorrow should reflect public concerns about privacy, transparency and accountability. Notably, healthcare organisations in the UK are increasingly seeking to involve the general public in data governance, with citizens juries and assemblies being pioneered around the country.49

Were we to outsource data clean up or some elements of the process, it would require us to carefully consider our own ethical frame of reference and regulations – and how these would extend to new trade partners. It would also potentially impact people’s perception of their ability to retain control over how their health data is used.

In recent weeks, the Prime Minister and officials setting out to negotiate our future relationship with the European Union have made clear the UK plans its own data protection regime post-Brexit. Jeni Tenison, CEO of the Open Data Institute, has indicated that she can see both challenges and opportunities in this which merit further discussion and a close watching brief.50 Future Care Capital has, for its part, called for additional protections to safeguard the UK’s healthcare data assets against this backdrop.51

**Conclusion**

For now, the UK remains an attractive country for the life sciences and advanced technologies like AI.52 However, to stay ahead it must develop a forward-thinking national strategy on how we harness the value of healthcare data – ahead of future trade negotiations beyond the EU. It is within the Government’s gift to seize this unique opportunity to unleash the potential of the UK’s unique healthcare data asset and increase the health and wealth of the nation.

If a major barrier to harnessing the value of healthcare data to develop new drugs and treatments is the lack of good quality and well-structured data,53 why would the Government not support the creation of a new data curation or advanced data services industry in the UK? Investment should flow from the Treasury toward left-behind communities, forging strong links between them and existing centres of excellence to help rebalance the UK economy. That could be a game-changing legacy for the new Government.

---

51 Annemarie Naylor, Briefing for House of Lords Debate on Protecting the NHS in Future Trade Deals (Future Care Capital, 2019).