



Healthcare systems must get fair value for their data

Stephen Bradley and colleagues call for action to ensure equitable returns for patients and taxpayers when public healthcare data are shared for commercial research

Stephen H Bradley,¹ Scott Hemphill,² Sarah Markham,³ Shivan Sivakumar⁴

The threats to patient privacy from sharing data have been well publicised. By comparison, the risks of healthcare systems missing out on the gains generated from healthcare data have been largely overlooked. As well as potential benefits from improved treatment or diagnosis, research using patient data can bring financial rewards—for example, by selling artificial intelligence (AI) products the data are used to develop. The indicative market value of the data held by England's NHS has been estimated at £5bn if it were to be sold for commercial purposes.¹

For cash strapped healthcare systems, stewardship of vast reserves of data presents opportunities for innovative collaborations with industry. However, they are not well placed to ensure fair returns. Although ethics committees examine the privacy implications of research using patient data, they do not generally consider protecting the value of data; nor are such committees well equipped to do so. Mechanisms need to be instituted to share revenues and access to new technologies that arise from private sector collaborations.

Government policy on collaborations

The UK government's strategy for the life sciences sector highlights the potential for the NHS to embark on partnerships with the business sector.^{2,3} Launched five years ago as a cornerstone of the government's long term economic programme, implementation of the strategy has fallen short of its ambitions, with data sharing processes remaining inconsistent and unclear.⁴

In England, a recent review into use of healthcare data for research commissioned by the secretary of state for health and social care (the Goldacre review) called for these processes to be streamlined and clarified.⁵ The report also advocates the creation of platforms through which healthcare data may be accessed and analysed. Such "trusted research environments" would ensure greater control over data by rationalising and regulating the types of information being accessed and preventing wholesale transfers of data. The report acknowledges that although trusted research environments can resolve privacy concerns, "there is a need for a frank public discussion about commercial use of NHS data."

Controversial collaborations

Several high profile transactions involving patient data have shown the need for accountability and transparency in terms of both the data shared and the value of what is received in return for patients and health systems. A collaboration between DeepMind (a subsidiary of Alphabet, owners of

Google) and the Royal Free NHS Foundation Trust used patient data to create an app to identify acute kidney injury.⁶ The project entailed sharing a wide range of healthcare data on 1.6 million patients and was subsequently deemed to have proceeded without an appropriate legal basis.^{7,8} A memorandum of understanding between both organisations set out the aspiration to establish a "broad ranging" partnership through which the trust would receive, besides bespoke software, "reputational gain" and a "place at the vanguard of developments in ... one of the most promising technologies in healthcare."⁸ As DeepMind is an AI company and the app simply implemented an existing NHS algorithm, concern has been expressed that the company's motivation may have been to acquire data for machine learning research.^{8,9}

Another DeepMind-NHS collaboration has used AI to predict progression to wet age-related macular degeneration through interpretation of optical coherence tomography scans—an application with substantial promise in establishing more consistent and efficient triage of patients in busy eye clinics.¹⁰⁻¹² DeepMind has made its AI model available to its NHS collaborator, Moorfield's Eye Hospital, but has not disclosed whether this access is time limited or under what terms it will offer this software to other NHS organisations or healthcare systems.¹¹

Controversial collaborations are not confined to individual trusts. NHS England agreed a contract with Amazon that offered the tech giant access to "all healthcare information" aside from patient records, apparently with nothing in return for the health service.¹³ Meanwhile, in the United States a partnership with the hospital chain Ascension that gave Google access to the detailed health data of millions of patients without their knowledge in return for data storage and software tools¹⁴ has been investigated by the Department of Health and Human Services' Office for Civil Rights.¹⁵

It is striking that in these cases the data sharing arrangements came to light only after investigations by journalists. The resulting discussion has tended to be framed in terms of consequences for patient confidentiality, chiming with a critique of a business model that relies on extracting, combining, and commodifying personal data, something that has been described as surveillance capitalism.^{16,17} The protections on healthcare data are more stringent than those for data which consumers have notionally consented to share. Nevertheless, even if these protections are followed, the lack of recourse to scrutinise the terms of data transactions, or even to discover that such agreements exist is troubling.^{6,18,19}

¹ Leeds Institute of Health Sciences, University of Leeds, Leeds, UK

² Sir Charles Gairdner Hospital, Perth, Australia

³ Visiting Researcher, Department of Biostatistics and Health Informatics, King's College London, London, UK

⁴ Department of Oncology, University of Oxford, Oxford, UK

Correspondence to: S Bradley
medsbra@leeds.ac.uk

Cite this as: *BMJ* 2022;377:e070876

<http://dx.doi.org/10.1136/bmj-2022-070876>

Published: 27 June 2022

The repeated attempts to institute routine sharing of English primary care data at scale show that it is not enough simply to demonstrate that such plans are not illegal.^{20,21} Public and professional acceptance requires that any use of data for means other than that originally intended must be in service of the public good without disproportionately benefiting other interests.²²

Obstacles to protecting value of healthcare data

Several factors may prevent the value of health data being realised in collaborations with commercial partners. Unlike commercial companies, many healthcare providers lack specialist expertise in commercial law and intellectual property, making them vulnerable to asymmetric agreements that benefit private sector collaborators.²³ Since agreements made with technology companies are often not made public, it is not possible to monitor whether they deliver proportionate value to taxpayers and health systems. The opacity of such arrangements has been likened to a “one way mirror” through which technology companies are able to analyse and profit

from patient data but the public cannot see how the data are being used and what, if anything, the healthcare system can expect to receive in return.^{8,24} Policy research has highlighted the need for greater accountability, and a recent public consultation emphasised that transparency is paramount throughout the data lifecycle.^{25,26}

Polling and qualitative research indicates broad support for sharing patient data with commercial partners provided that it delivers demonstrable public benefit and does not disproportionately reward private interests.²⁴ But defining what constitutes sufficient public benefit is challenging, and it will vary from case to case. However, we contend that satisfying this test requires more than the development of proprietary technology that could be used to improve patient care. When technology has been developed using patient data or co-produced using health system resources, the resulting value should be returned through proportionate mechanisms such as cost-free access to the technology or a revenue share for the health service (table 1).

Table 1 | Models of value sharing between health systems and private sector. Adapted from Ghafur et al²⁷

| Agreement | Description | Potential concerns | Example |
|---------------------------------------|--|--|---|
| No value sharing | Health system shares data for free | Health system receives no share of value of data | Amazon-NHS ¹³ |
| Free or discounted products | The product developed is provided to the health system organisation or to the whole of the health system for free or at a discount (for a defined or unlimited period) | No value captured from non-UK income. If the product is discounted or free only for one organisation, other organisations within the health service will have to pay | DeepMind-Moorfields ²⁸ |
| Royalty or revenue share | Health system receives a royalty or a portion of the revenue from products developed using its data | Health system does not share ownership of the products being developed | Sensyne Health-Oxford University Hospitals NHS Trust (also includes equity share) ²⁹ |
| Profit share | Health system receives a royalty or a portion of the profits from products developed using its data | Health system does not share ownership of products being developed. Value captured will depend on the profitability of the company, rather than product revenues | |
| Intellectual property ownership share | The health system receives partial ownership of the intellectual property generated | High cost and complicated arrangement | |
| Equity share | Health system receives a share of the equity of the company developing solutions from the data | Unappealing to established companies | Sensyne Health-Oxford University Hospitals ²⁹ |
| Fee for access | Health system receives a one-off payment in exchange for access to the data | Depending on the pricing, health systems could get limited value. May penalise smaller companies that have less funding | Hospital Episode Statistics data from NHS Digital ³⁰ |

How should the value of data be protected?

Failing to protect health data as valuable public assets risks making taxpayers pay both to develop and to use novel technologies. This costly model of discovery is now entrenched elsewhere in medicine. Notably, the pharmaceutical industry deploys assertive pricing strategies for drugs developed from publicly funded research, and scientific publishers obtain research and editorial services from academics before selling this back to the publicly funded institutions that employ those academics.³¹⁻³³ Much focus has been dedicated to highlighting and exploring potential remedies to these problems, including greater regulation and more assertive involvement of the state in innovation.^{34,35}

The issue of safeguarding the value of healthcare data therefore resonates with wider concerns about profit from public assets being diverted to private interests. The nascent status of technologies such as AI presents an opportunity to formulate regulations and norms to protect value. But these issues are particularly complex with respect to healthcare data, not least because multiple

organisations and individuals may be said to have contributed to its creation and curation.³⁶ We also currently lack frameworks to determine acceptable remuneration in cash, in kind (eg, access to products), or percentage stake in any resulting profits, that should be returned to health systems.

Formulating how health systems should share rewards resulting from collaborations using patient data is therefore far from straightforward. Innovations that have been created entirely independently and are simply validated in a healthcare setting could be judged to be analogous to devices or drugs, with the manufacturer retaining the prerogative to negotiate on pricing without reference to the contributions made by the health service in validating the technology. But, when health systems contribute substantial resources to evaluating new technologies, such as with the NHS Grail study which aims to detect cancer in asymptomatic patients using a novel blood test,³⁷ or when AI algorithms are validated or improved based on performance, there may be a rational claim to some form of reimbursement.

While most ventures using healthcare data will never prove profitable, there should not be a presumption that it is acceptable to harvest patient data for any innovation that might be commercially or clinically successful. Nor should healthcare systems be expected to underwrite the costs and risks of collaboration in the name of innovation. But there is a strong case that healthcare services and the public research institutes should share the benefit from any collaborations that do generate revenues because they will also bear the costs of efforts that are unsuccessful.³⁸

This might be achieved through arrangements such as healthcare systems, or the state, taking an equity share in collaborative ventures.^{27 38} Alternative means of sharing value more directly with patients who contribute data, rather than healthcare systems or the state, have been envisaged, including royalty payments to individual patients and creation of independent community development funds for relevant populations.^{16 39 40}

As the Goldacre review highlighted, adequate consultation on these issues by governments and health systems is overdue.⁵ Creation of frameworks to guide expectations of value sharing requires expertise and perspectives of ethicists, intellectual property specialists and healthcare technology specialists, industry representatives, healthcare staff, patients, and the public. Such consultations could include public deliberative procedures such as citizen's assemblies. **Box 1** lists some questions that could be considered. Experience, as well as numerous reports and consultations, have emphasised that giving the public a say in how people's data are used is crucial to establishing and maintaining trust, which is essential for fruitful collaboration.^{5 25 41} Failure to invest the time and resources in adequate public and professional consultation to create a robust foundation for private-public collaboration using health data is likely to lead to greater cost and delay to innovation in the long term.

Box 1: Questions to address on collaborations using healthcare data

- What safeguards are necessary to fund costs for health services from collaborations, such as preparing data, to ensure resources are not diverted from delivering patient care?
- For the minority of ventures that return revenues, should these be returned to local health service organisations, to the central health service, or to the nation's treasury?
- What kinds of partnership models are suitable for small and medium sized commercial partners versus those that are appropriate for larger companies?
- Should preferential terms apply for domestic companies, as opposed to overseas firms, as a means to foster wider benefits to society and the economy, such as employment and taxation?
- Should organisations which act as subcontractors to health services, such as general practices in the UK, be permitted to negotiate value sharing collaborations independently?

Healthcare systems need to invest in staff with expertise in negotiating intellectual property agreements to support those working for healthcare services who wish to use data in pursuit of innovation. Such agreements should be made publicly available. When centralised expertise is available to support health systems to collaborate with industry, organisations within those systems should be incentivised, or possibly even required, to accept that support, rather than striking deals on their own.⁴² Central scrutiny bodies, such as the National Audit Office in the UK or the Government Accountability Office in the US could be responsible for ensuring adequate value is returned to the public realm through commercial partnerships.

Collaborations based around applying technologies such as AI to healthcare data promise to unlock new discoveries with both commercial and clinical value. But the public has a vital stake in determining how the value that results from such products is distributed and whether it is reasonable for such collaborations to proceed at all. Neglecting these legal and ethical frontiers in pursuit of innovation risks ceding valuable assets to private interests and could prove a costly legacy for patients and taxpayers.

Key messages

- Data from healthcare systems hold value for improving healthcare delivery and in the development of commercially successful products through private sector collaborations
- As well as safeguarding privacy, data sharing agreements must ensure fair benefit for health systems and the public
- Lack of commercial expertise and transparency risks health systems being disadvantaged in agreements
- Health systems and governments must establish terms for sharing data informed by extensive public, professional, and expert consultation

Contributors and sources: This article was conceived by SHB and SS, who had become aware of the issues of data sharing and the need to recoup value for patients and taxpayers through their familiarity with innovations in cancer diagnosis. Additional background information on the policy context was obtained through discussion with colleagues who have expertise in health informatics and review of relevant policy reports. SM and SH provided input to revisions of the manuscript.

Public and patient involvement: Two members of use MY data's advisory group provided feedback on the draft manuscript. One member (SM, who is also a member of BMJ's patient panel) then became a co-author of the article.

Competing interests: We have read and understood BMJ policy on declaration of interests and declare the following: SHB has submitted a grant application for a project which plans to evaluate the performance of AI systems in diagnosing lung cancer. He is clinical lead for cancer for NHS Leeds clinical commissioning group. He receives PhD funding from CanTest collaborative (Cancer Research UK C8640/A23385) and was a member of the executive committee of the Fabian Society, a political think tank affiliated with the Labour Party (unpaid). The publication costs of a collection of essays on health inequalities which he co-edited for the Fabian Society were funded by the Association of the British Pharmaceutical Industry and Lloyds Pharmacies. He has received funding from the Mason Medical Foundation for a study on lung cancer diagnosis.

Provenance and peer review: Not commissioned; externally peer reviewed.

We thank Saira Ghafur and Jessica Butler for feedback.

- 1 Ernst and Young. Realising the value of health care data: a framework for the future. 2019. https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/life-sciences/life-sciences-pdfs/ey-value-of-health-care-data-v20-final.pdf
- 2 HM Government. Life sciences: industrial strategy. 2017. <https://www.gov.uk/government/publications/life-sciences-industrial-strategy>
- 3 Gibbons M, Limoges C, Nowotny H, Schwartzman S, Scott P, Trow M. *The new production of knowledge: the dynamics of science and research in contemporary societies*. Sage, 1994.
- 4 House of Lords Science and Technology Committee. First report of session 2017-19. Life sciences industrial strategy. Who's driving the bus? 2018;26. <https://publications.parliament.uk/pa/ld201719/ldselect/ldscctech/115/115.pdf>
- 5 Goldacre B, Morley J. Better, broader, safer: using health data for research and analysis. Department of Health and Social Care, 2022. <https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis>
- 6 Hodgson H. Revealed: Google AI has access to huge haul of NHS patient data. *New Sci* 2016;2016.
- 7 Iacobucci G. Patient data were shared with Google on an "inappropriate legal basis," says NHS data guardian. *BMJ* 2017;357. doi: 10.1136/bmj.j2439 pmid: 28522583
- 8 Powles J, Hodson H. Google DeepMind and healthcare in an age of algorithms. *Health Technol (Berl)* 2017;7:67. PubMed doi: 10.1007/s12553-017-0179-1
- 9 NHS England. Algorithm for detecting acute kidney injury (AKI) based on serum creatinine changes with time. 2014. <https://www.england.nhs.uk/wp-content/uploads/2014/06/psa-aki-alg.pdf>
- 10 Yim J, Chopra R, Spitz T, et al. Predicting conversion to wet age-related macular degeneration using deep learning. *Nat Med* 2020;26:9. doi: 10.1038/s41591-020-0867-7 pmid: 32424211
- 11 Suleyman M. A major milestone for the treatment of eye disease. DeepMind, 2018. <https://deepmind.com/blog/article/moorfields-major-milestone>
- 12 Powell S. Beautiful mind: what AI means for the future of optometry. 2019. <https://www.aop.org.uk/ot/science-and-vision/technology/2019/06/28/beautiful-mind>

- 13 Mahase E. Government hands Amazon free access to NHS information. *BMJ* 2019;367. doi: 10.1136/bmj.l6901 pmid: 31818911
- 14 Shaikat T. Our partnership with Ascension. Inside Google Cloud, 11 Nov 2019. loud.google.com/blog/topics/inside-google-cloud/our-partnership-with-ascension
- 15 Copeland R, Needleman SE. Google's 'project nightingale' triggers federal inquiry. *Wall Street Journal* 2019 Nov 12. <https://www.wsj.com/articles/behind-googles-project-nightingale-a-health-data-gold-mine-of-50-million-patients-11573571867>
- 16 McCoy MS, Joffe S, Emanuel EJ. Sharing patient data without exploiting patients. *JAMA* 2020;323-6. doi: 10.1001/jama.2019.22354 pmid: 31944215
- 17 Zuboff S. Surveillance capitalism and the challenge of collective action. *New Labor Forum* 2019;28-29doi: 10.1177/1095796018819461
- 18 Das S, Gregory A. Amazon ready to cash in on free access to NHS data. *Sunday Times* 2019 Dec 8. <https://www.thetimes.co.uk/article/amazon-ready-to-cash-in-on-free-access-to-nhs-data-bb2p52n5m#:~:text=Amazon%20has%20been%20handed%20the,a%20pen-ny%20to%20the%20UK>
- 19 McDonald AM, Cranor LF. The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society* 2009. <https://lorrie.cranor.org/pubs/readingPolicyCost-authorDraft.pdf>
- 20 de Zulueta P. Confidentiality, privacy, and general practice: GDPR and the brave new world of 'big data'. *Br J Gen Pract* 2021;71-1. doi: 10.3399/bjgp21X717017 pmid: 34446416
- 21 van Staa T-P, Goldacre B, Buchan I, Smeeth L. Big health data: the need to earn public trust. *BMJ* 2016;354-. doi: 10.1136/bmj.i3636 pmid: 27418128
- 22 Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care.data ran into trouble. *J Med Ethics* 2015;41-9. doi: 10.1136/medethics-2014-102374 pmid: 25617016
- 23 O'Carroll M, Kimbro S. In: Katz DM, Dolin R, Bonmarito MJ, eds. *Legal Informatics*. Cambridge University Press, 2021.
- 24 Wellcome Trust. The one-way mirror. Public attitudes to commercial access to health data, 2016. <https://wellcome.org/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf>
- 25 Ada Lovelace Institute. Algorithmic accountability for the public sector: learning from the first wave of policy implementation. 2021. <https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector/>
- 26 National Data Guardian. Putting good into practice: a public dialogue on making public benefit assessments when using health and care data. 2021. <https://www.gov.uk/government/publications/putting-good-into-practice-a-public-dialogue-on-making-public-benefit-assessments-when-using-health-and-care-data>
- 27 Ghafur S, Fontana G, Halligan J, O'Shaughnessy J, Darzi A. *NHS data: maximising its impact on the health and wealth of the United Kingdom*. Imperial College London, 2020.
- 28 Moorfields Eye Hospital NHS Foundation Trust. DeepMind Health research partnership. Q&A. 2019. <https://www.moorfields.nhs.uk/faq/deepmind-health-qa>
- 29 Sensyne Health signs strategic research agreement with Oxford University Hospitals NHS Foundation Trust. Press release, 7 Aug 2018. <https://www.sensynehealth.com/newsroom/sensyne-health-signs-strategic-research-agreement-with-oxford-university-hospitals-nhs-foundation-trust>
- 30 NHS Digital. Data access request service charges from 2020/2021. <https://digital.nhs.uk/services/data-access-request-service-dars/data-access-request-service-dars-charges>
- 31 Mazzucato M. High cost of new drugs. *BMJ* 2016;354-. doi: 10.1136/bmj.i4136 pmid: 27815284
- 32 Aczel B, Szaszi B, Holcombe AO. A billion-dollar donation: estimating the cost of researchers' time spent on peer review. *Res Integr Peer Rev* 2021;6-. doi: 10.1186/s41073-021-00118-2 pmid: 34776003
- 33 Buranyi S. Is the staggeringly profitable business of scientific publishing bad for science? *Guardian* 2017 Jun 27. <https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science>
- 34 Darrow JJ, Light DW. Beyond the high prices of prescription drugs: a framework to assess costs, resource allocation, and public funding. *Health Aff (Millwood)* 2021;40-8. doi: 10.1377/hlthaff.2020.00328 pmid: 33523733
- 35 Mazzucato M, Chow H, Fitzpatrick S, et al. *The people's prescription: re-imagining health innovation to deliver public value*. University College London Institute for Innovation and Public Purpose, 2018.
- 36 Alami H, Lehoux P, Auclair Y, et al. Artificial intelligence and health technology assessment: anticipating a new level of complexity. *J Med Internet Res* 2020;22:e17707. doi: 10.2196/17707 pmid: 32406850
- 37 NHS England. NHS to pilot potentially revolutionary blood test that detects more than 50 cancers: 2020. <https://www.england.nhs.uk/2020/11/nhs-to-pilot-potentially-revolutionary-blood-test/>
- 38 Laplane A, Mazzucato M. *Socialising the risks and rewards of public investments: Economic, policy and legal issues*. University College London Institute for Innovation and Public Purpose, 2019.
- 39 Cassel C, Bindman A. Risk, benefit, and fairness in a big data world. *JAMA* 2019;322-6. doi: 10.1001/jama.2019.9523 pmid: 31287512
- 40 Winickoff DE, Winickoff RN. The charitable trust as a model for genomic biobanks. *N Engl J Med* 2003;349-4. doi: 10.1056/NEJMs030036 pmid: 13679534
- 41 One London, Ipsos, King's Fund. Public deliberation in the use of health and care data. 2020. <https://www.onelondononline/wp-content/uploads/2020/06/Public-deliberation-in-the-use-of-health-and-care-data.pdf>
- 42 NHS England. The Centre for Improving Data Collaboration. <https://www.nhs.uk/key-tools-and-info/centre-improving-data-collaboration/centre-improving-data-collaboration-what-we-do/>