KONP Data Working Group's response to Department of Health & Social Care's *Data saves lives: reshaping health and social care with data*[1]

Part One:

Overview

The strategy document *Data saves lives: reshaping health and social care with data* sets out how the Department of Health and Social Care wants data to be used to bring benefits to all parts of health and social care, from patients and care users to front line staff and pioneers driving research. It's supported by a number of commitments, including investing in secure data environments and giving people better access to their data through shared care records and the NHS App.
The document emphasises the government's wish to create the public trust necessary to use our personal health data to achieve their stated ambitions: to create innovations necessary for 'sustainable' health systems[2], and to drive growth in the UK economy[3] in order to power our recovery, including by becoming a global AI superpower[4]

Our response focuses on what the proposals say about the development of 'trust', providing access to data for 'wider' purposes, the power of the new legislation, the development of the data analytics ecosystem (including the new Federated Data Platform) and the use of 'open data' to bring private provision into health systems through a 'market place' for innovation and consumption.

Our key points:

- Public trust in the use of our personal health data is essential but needs to be present not only for the NHS, but also in any use by the government and their commercial 'partners'
- Experience of commercial 'partners' in the NHS has not imbued trust
- The strategy's call for data altruism – in effect a citizen's duty to share their data - places unfair burdens on individuals to solve difficulties not of their making
- The value of NHS data can not be expressed nor understood purely in fiscal terms
- The increased access to data allowed under Control Of Patient Information regulations during the COVID pandemic must not be used routinely, as proposed: this would be to the detriment of our rights
- The powers of NHS Digital must remain separate from NHSE to avoid conflicts of interest and misuse of data when it comes to plans to restrict resource allocation; and also to avoid pressures on NHSE to share data with other public sector bodies regarding, for example, welfare benefits or immigration
- No personal health data should leave secure data environments

Part Two:

Trust

---

[1] https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data/data-saves-lives-reshaping-health-and-social-care-with-data

[2] https://www.longtermplan.nhs.uk/

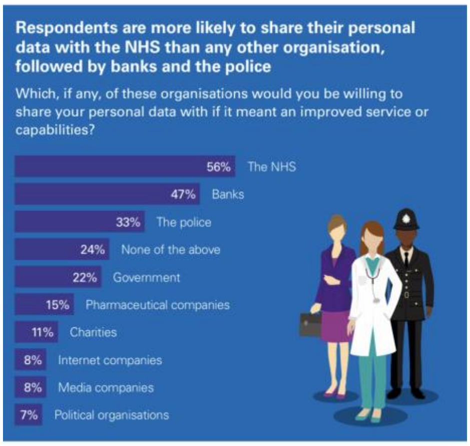
[3] https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy

[4] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1020402/National_AI_Strategy_-_PDF_version.pdf

There will be innovations that could very well improve our healthcare diagnoses, treatment and prevention. However, given the huge involvement of ‘private partners’ in the IT infrastructure of our health systems[5], and in the management and analysis of our data[6], the price of any potential innovations remains unclear. The Government rightly believes that we will not allow our data to be used unless we ‘trust’ them. They might also add that we need to ‘trust’ their ‘partners’ - a usual euphemism for industry or private companies.

Trust has certainly not been increased by the profiteering of private healthcare organisations in the early phases of the Covid pandemic [7,8,9] and neither is there much to reassure us in the history of involvement of private enterprise in the NHS and social care.[10,11,12]

Back in 2018, the global consultancy firm KPMG[13] were pushing government to concentrate on building trust in the public for the sharing of their personal health data if they wished to ‘win the AI race’: a survey by KPMG showed a complete lack of trust in private companies -see their chart - and it was necessary to ‘turn this around’ if such companies were to be involved as ‘partners’ and allowed to profit (the only reason that they would be involved!) from a public asset - the NHS database of our personal health data.



KPMG survey on trust

Now, with the arrival of major healthcare corporations such as Centene[14], UnitedHealth[15] and Palantir from the US, the public are becoming aware of issues such as malpractice, major lawsuits and fights in court, fixations on profit margins to the detriment of care[16] and links of some corporations to government security services[17].

It seems that government wishes to hide the use of our data by their commercial ‘partners’ behind: i) trumpeting their efforts to make data secure and safe; ii) banking on our trust in the NHS and for

---

[5] https://www.england.nhs.uk/hssf/supplier-lists/

[6] https://www.england.nhs.uk/hssf/supplier-lists/#informatics-analytics-and-digital-tools-for-population-health-business-and-clinical-intelligence

[7] https://www.theguardian.com/world/2022/mar/30/profiteering-over-covid-ppe-disgraceful-says-uk-government-adviser

[8] https://www.theguardian.com/business/2022/feb/08/pfizer-covid-vaccine-pill-profits-sales

[9] https://www.bmj.com/content/374/bmj.n1750

[10] https://chpi.org.uk/resources/pfi-companies-profited-nhs/

[11] https://chpi.org.uk/wp-content/uploads/2021/09/CHPI-For-Whose-Benefit_.pdf

[12] https://chpi.org.uk/papers/reports/plugging-the-leaks-in-the-uk-care-home-industry/

[13] https://assets.kpmg/content/dam/kpmg/uk/pdf/2018/09/how-the-uk-can-win-the-artificial-intelligence-ai-race.pdf

[14] https://violationtracker.goodjobsfirst.org/parent/centene

[15] https://violationtracker.goodjobsfirst.org/parent/unitedhealth-group

[16] https://keepournhspublic.com/bbcs-panorama-undercover-britains-biggest-gp-chain-showed-us-why-big-business-has-no-place-in-our-gp-services/

[17] https://www.theguardian.com/commentisfree/2022/jun/14/palantir-the-all-seeing-us-tech-company-could-soon-have-the-data-of-millions-of-nhs-patients-my-response-yikes

publicly funded researchers, iii) and, most hypocritically, implying that individuals have a duty to allow their data to be used to improve health for us all (altruism).

For example, Tim Ferris, Director of NHSE Transformation[18] has spoken about the [power] of systems which can be used to personalise altruism - 'where you start thinking about the ethics of your commitment to the community….my opting out makes it more difficult for the community to cure cancer….[by not opting out] we are doing our duty as citizens'.

When looking at how else large, complete data sets are expected to help health and social care, arguments could also be framed in a specific way so as to address individual health and care systems: 'to make it work' and more specifically 'to maximise its value' and 'to reduce its inequalities'. Localising responsibility in this way places burdens on individual members to solve issues not of their creation, and can act to play down the responsibilities that the state should take up, such as to provide 'adequate' funding or to legislate to reduce inequalities[19].

Data saves lives proposes five ways of demonstrating that the health and care system will be a trustworthy data custodian, namely by:

1. Keeping data safe and secure.
2. Being open about how data is used.
3. **Ensuring fair terms from data partnerships.**
4. Giving the public a bigger say in how data is used.
5. Improving the public's access to their own data.

Nothing is said about who will use the data, and there seems to be unwillingness to mention 'commercial use', which is a poor start to transparency, a prerequisite for trust.

In contrast, Nicola Byrne the National Data Guardian (NDG) in her Oral evidence to the Science and Technology Committee stated that:

> 'Public benefit must come above commercial profit and be the primary consideration. People want to see that safeguards are in place and that there are meaningful sanctions for any improper use'.

She expressed two further concerns:

> 'I would not want to see any financial contribution [from business] mean that those ethical obligations were somehow paid off and were no longer to be considered'

and

> 'The value of (NHS) data is not to be expressed or understood purely in fiscal terms in the short to medium term. People understand that the value of this data is, potentially, knowledge generation in years way beyond where we are now….. The NHS has to be extremely careful about being locked into any commercial contracts that might give exclusivity to that knowledge in future and lock it away from us'.

Commercial companies would need to be completely transparent about their use of data, their own data processing, and the reasons for its use currently and in the future. This, as campaigners know, is in conflict with the legal realities of protecting intellectual property rights and commercial practices that companies will fight for 'tooth and nail'- but that is what needs to be done and may prove to be impossible.

Providing access to data for wider purposes

---

[18] Science and Technology Committee. Oral Evidence: The right to privacy: digital data, HC 1000. House of Commons Wednesday 20 April 2022.

[19] https://localdemocracyandhealth.files.wordpress.com/2015/06/neoliberalism-with-a-community-face-2.pdf

Since the outbreak of Covid-19, existing protections for our health data have been waived under the Control of Patients Information (COPI).[20] COPI regulations, introduced in 2002 without any Regulatory Impact Assessment, give the Secretary of State (SoS) temporary powers to allow the processing of confidential information during a public health emergency. The Regulation's explanatory notes state:

> 'In general the Regulations enable the flow of information and impose no obligations. Where obligations are imposed, they are imposed primarily on those performing functions for public authorities and so any burden imposed on business is considered negligible'.

In March 2020, Regulation 3 of the Act was invoked, requiring NHS Digital to disseminate confidential patient information to those employed or engaged for the purposes of the health service in response to the Covid pandemic. The measure was initially to last for six months but extended until June 2022. However, the Government is now arguing that non-emergency regulations are insufficient to support the data requirements for the 'proper management and development' of the new health and care systems. They insist that the access to, and use of, confidential patient information should be increased to become more 'proportionate and appropriate' with minimal delay, while reassuring us that they will maintain confidence in how our data is protected.

The Government's aim is therefore to amend and use the 2002 COPI regulations to provide the kind of access to data it allows even beyond a state of emergency. Their stated aim is that 'researchers will be able to safely and easily access data to provide innovative solutions to health and care issues for the benefit of all'. This of course includes private companies, and charitable organisations with private partners/sponsors as major 'innovators', but they do not mention this.

Using the Health and Care Act 2022

Powers in the Health and Care Act 2022 will be used to ensure completeness of the NHS database, with added databases from social care and private provision, in a standardised format 'ripe and ready' for anyone who accesses it, for whatever purpose. The powers will: 'require' that anonymous data flows freely across all partners in the health and care systems; allow NHS Digital to access any data 'connected with' health care and promotion including from private providers; and mandate standards for processing data to render it meaningful and usable for whatever purpose (this will include private health and adult social services where they relate to the provision of NHS services).

Part 3 of the Health and Care Act 2022[21] also includes provisions that allow the SoS to make regulations to transfer functions between particular named NHS bodies, and there is the intention to use this power to transfer the functions of NHS Digital (until now relatively independent) as far as possible to NHS England.

However, as Kingsley Manning (the ex-chair of NHS Digital) has stated[22], this merger is a "grave error", especially given the Government's intention to weaken constraints on the use of patient data. He points out that NHSE's priorities are often in conflict with the interests of individual patients or patient groups, while NHSE's decisions on resource allocation are increasingly data driven, and made at the micro level.

Moreover, with the enhanced power of AI and access to ever more data sets, including those beyond the health care system, NHSE will have the ability to aggregate a patient's data from all their interactions with the NHS. This information, coupled with opaque algorithms, can then be used as a basis as to whether or not to allow treatment – for example a particular drug therapy - without any requirement to explain to a patient how this decision has been made or what technologies have been applied.

---

[20] https://www.legislation.gov.uk/uksi/2002/1438/contents/made

[21] https://www.legislation.gov.uk/ukpga/2022/31/part/3/enacted

[22] https://www.bmj.com/content/376/bmj.o361/rr-0

In addition, NHSE could be pressed to share and pool a patient's data with other public sector bodies to inform a wider range of decisions, for example, regarding criminal cases, welfare benefits, social care, employment, immigration and so on.

Improving the data analytics ecosystem

Government believes that the NHS 'suffers from siloed data from multiple sources…. and a lack of consistent and timely access….which will limit integrated care systems in achieving their goals'. Their answer to this is the development of Secure Data Environments (SDEs) - a subset of which will be Trusted Research Environments (TREs).

Back in 2021,[23] and later in the 2022 Goldacre Review[24], TREs seemed to be the favoured solution. However, it seems that government feels that TREs may not make the most out of secure environments, and that it must continue to examine how these environments operate and are governed to ensure SDE's are fit for the future. They are therefore bringing together their 'partners' across the system to set out the role of secure data environments in the health and care system, the standards they must meet and policies to govern their use.[25] This will include 'enhancing the exchange of skills and knowledge' to enable closer working between groups.

With SDEs, which are expected to be the default solution by December 2022 (an incredibly short time-frame), data can be accessed securely in a virtual setting, while analysis – we are assured - takes place within a secure online platform, rather than being shared and distributed, and access to data is only granted to authorised researchers in a controlled and recorded manner. However, information *can* leave the server:

> 'Users' interactions with the data will be recorded and monitored, and the information they can extract will be assessed and with personal identifiers removed. *As no data that can be linked to an individual leaves the server*, and all access to the data and analysis is monitored….we will greatly reduce the risk of data breaches or other misuse.' (Our italics)

Note that they do not state that extracted data will be anonymised data. If it simply has 'personal identifiers removed' this could be pseudonymised data: "the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organizational measures to ensure non-attribution to an identified or identifiable individual" (Article 4(3b) of General Data Protection Regulation or GDPR). Such data can be de-anonymised - a reason given by Ben Goldacre[26] for withdrawing his consent from the recent GP Data for Planning and Research programme (GPDPR), set up to extract data from GP records (commonly known as 'GP-data grab' and temporarily paused due to the number of patients 'opting out').

Subject to HM Treasury approval', the next attempt to reinvigorate the GPDPR is going to be a 'flagship' example of a service where data will only be accessible via an SDE- *this implies other services may have data accessible in other ways, and still does not mean that data cannot be taken away*.

---

[23] https://www.hdruk.ac.uk/news/new-principles-published-to-improve-public-confidence-in-access-and-use-of-data-for-health-research-through-trusted-research-environments/

[24] https://www.gov.uk/government/publications/better-broader-safer-using-health-data-for-research-and-analysis

[25] https://www.nhsx.nhs.uk/blogs/joining-up-the-dots- driving-innovation-research-and-planning-through-trusted-research-environments/ link given in document but not available as moving to the NHS Transformation directorate

[26] House of Commons Science and Technology Committee. Oral evidence: The right to privacy: digital data, HC 97, 11 May 2022

When, in 2023, NHS England assumes responsibility for NHS Digital's functions (subject to parliamentary processes), it claims that a statutory safe haven for health and care data will be created within NHSE, and its use of data, including how it shares data externally, will be subject to independent scrutiny. This *apparently* will ensure transparency and the continuation of the high standards of data security as exemplified by NHS Digital. The myth of such standards has been shattered by independent investigations, as shown, for example by *they sold it anyway's*[27] report on where patients' hospital data is sold; the Financial Times' investigations[28] uncovering multiple sales of data to private companies that were unknown to the public; and the NHS Digital audit reported in the BMJ, which found that hundreds of health data breaches were left unpunished.[29]

A Federated Data Platform

Having apparently learned lessons and developed expertise about data management from their COVID-19 response, government feel the need for further improvement by developing a Federated Data Platform (FDP)[30] which will be a national system of connected platforms, placed in, and ultimately determined by, individual NHS organisations.

Apparently, this technology infrastructure will:

- streamline operational and performance data collection from providers
- facilitate rapid development and sharing of analytics across integrated care systems and support operational improvement
- maintain integrated care system data control, and
- enable appropriate provider data to be accessible and visible to local, regional and national bodies for improved operational planning

The ambition is that every Trust and integrated care system will have their own platform that protects data to the highest standards of privacy and security, in accordance with the SDE requirements, which can also interact with regional and national platforms when they need to fulfil specific, predetermined use cases. The front-runner for the contract for this platform is Palantir, a US corporation that has many critics.[31] The FDP with its multiple platforms (there are 209 Trusts in England not including 10 Ambulance Trusts[32] and 42 ICSs) moving data all over the place and determining data usage represents almost the opposite to what the Goldacre Review[23] has called for in order to engender trust - i.e. as small a number of TREs as possible to manage all personal healthcare data, to avoid duplication of effort and to maintain strict control. It seems that in addition to the 251 platforms in the FDP, there will be a few TREs at national and regional level to support research.[33]

Working in the open and open data[34] (quite different)

It makes some sense to wish to 'work in the open' - for a publicly funded service to want to be transparent to its paymasters and share findings across the system (see the strategy's commitments 411 and 601). Such a way of working will include sharing technical skills and domain

---

[27] https://theysolditanyway.com

[28] https://www.ft.com/content/6f9f6f1f-e2d1-4646-b5ec-7d704e45149e

[29] https://doi.org/10.1136/bmj.o1126

[30] https://www.digitalhealth.net/2022/04/nhs-england-federated-data-platform/

[31] https://www.opendemocracy.net/en/opendemocracyuk/uk-health-department-ends-data-deal-with-spy-tech-company-palantir/

[32] https://www.kingsfund.org/audio-video/key-facts-figures-nhs

[33] The Data for Research and Development Programme (https://www.gov.uk/government/news/260-million-to-boost-healthcare-research-and-manufacturing) is investing up to £200 million in NHS data infrastructure (subject to HM Treasury approval) to make research-ready data available to researchers in a streamlined, secure and privacy-protected way, through trusted research environments by December 2023.

[34] Data is open if anyone is free to use, reuse and redistribute it-subject only at most to the requirement to attribute and share-alike-opendefinition.org

knowledge through sites like *Cross-validated*[35] and *Stack Overflow*[36], while sharing code and methodology through platforms like *GitHub*[37] in order to build high-quality analytics throughout the system. But even here there are attractions for business, such as easing introduction of new products into the health systems and ownership of the sites used for analysis development[38].

In contrast, the wish for Open Data follows through on a Memorandum of Understanding signed between the US and UK in 2018.[39,40] Open data is data that is accessible at little or no cost, exploitable, editable and shared by anyone for any purpose, even commercially. It is championed by NHS Digital[41] to improve outcomes, productivity, efficiency, choice and accountability in our public services while encouraging economic growth by 'supplying tools for public commissioners and providers of services'. Such tools can be seen in its 'Innovation Accelerator,'[42] an NHSE initiative delivered in partnership with the country's 15 Academic Health Science Networks (AHSNs) to help innovations developed by businesses to be spread at 'pace and scale' throughout the NHS.[43]

Significantly, transglobal business consultants McKinsey[44] think that the government has a critical role in transforming every sector of the economy in order to unleash more than $3 trillion in global economic potential from the use of 'open data'. McKinsey's advice to maximise the health market and generate maximum profits is for our government to: create a thriving ecosystem of data users, coders, and application developers; advertise open-data availability[45]; protect organisations from risk such as confidentiality, liability, and intellectual property[46]; participate in setting technical standards that can significantly increase and scale- up the financial benefits that open data can bring, while ostensibly protecting individuals from the risks of open data, concerning privacy, security, and personal safety.

No problem there then, this is exactly what our government is doing!

What is the alternative?

We argue that in a fully public service, in order to achieve a world class data system:

- Government should run competitive tendering from public institutions such as university departments and Institutes to develop the necessary skills base required to build and run an in-house data system for the NHS.

---

[35] https://machinelearningmastery.com/k-fold-cross-validation/

[36] https://stackoverflow.com

[37] https://github.com/

[38] In 2021, Prosus, an investment subsidiary of controversial South African media company Naspers, acquired Stack Overflow for US$1.8 billion; in 2018 Microsoft acquired GitHub for US$7.5 billion. It is commonly used to host open-source projects and as of June 2022, it reports having over 83 million developers and more than 200 million repositories (including at least 28 million public repositories)

[39] https://www.healthit.gov/sites/default/files/mou_v5_signed.pdf

[40] https://www.healthit.gov/sites/default/files/us_uk_standards_april2016.pdf

[41] https://digital.nhs.uk/services/supporting-open-data-and-transparency

[42] https://nhsaccelerator.com/innovations/page/3/

[43] https://www.digitalhealth.net/2022/04/nhs-innovation-accelerator-cohort-2022/

[44] https://www.mckinsey.com/industries/public-and-social-sector/our-insights/how-government-can-promote-open-data

[45] https://data.gov.uk/search?filters%5Btopic%5D=Health

[46] https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation

- Required resources for IT, such as Parallel Computing[47] data storage[48] or health software development[49] could also be further developed in public universities as well as tendered for by government for the NHS in the same way as surgical equipment and medication. There should be no Private Finance Initiatives nor other long-term dependency on private corporations, while specific skills transfer should be part of any contract.

- There should be ‘partnership’ working between the NHS and the relevant public institutions - access to NHS personal health data would be invaluable for public research and an advantage for grant achievement and international academic joint working for larger projects, where world leading universities and public institutes share resources, skills and the costs of joint research and development for the public good.

- Innovations produced between academic institutions and the NHS would be used in services, patented by government, and shared to other countries, perhaps on a variable cost basis depending on the receiver’s resources, for use in their health services. Money made should fund further work and improve patient care or even be fed back to the exchequer.

- Such public investment would require considerable initial investment but would avoid the profit margins and long term (possibly forever) subscription and update payments required by private providers.

July 2022

---

[47] https://www.epcc.ed.ac.uk

[48] https://www.qub.ac.uk/News/Allnews/2019/QueensandUniversityofGlasgowreceive4milliontotraindataexperts.html

[49] https://www.ed.ac.uk/bayes/about-us/education/workforce-development/courses/introduction-to-software-development-in-health