



INSIGHT | DATA SHARING DURING CORONAVIRUS

NHS Covid-19 Data Store and NHS National Data Platform

Summary of a private roundtable

Gavin Freeguard and Paul Shepley

Introduction

This short paper summarises a roundtable discussion held in summer 2022 about the NHS Covid-19 Data Store and NHS National Data Platform, which NHS England launched towards the beginning of the Covid pandemic. The roundtable brought together public servants, private companies and others involved in the project and was held under the Chatham House Rule: within this paper nothing anyone said is attributed to them or their organisation, unless they have asked for it to be. The discussion does not represent the views of the Institute for Government.

The roundtable forms part of a wider piece of Institute for Government research looking at government data sharing during the pandemic. The project takes six case study areas and uses a roundtable on each to explore what worked well, what could have worked better and what lessons government should learn for the future. Papers on each of the roundtables will be followed by a short synthesis report bringing together common themes.

Overview of the NHS Covid-19 Data Store and NHS National Data Platform

To respond to the pandemic, the government commissioned NHS England and NHS Improvement* and NHSX** (which merged with NHS Digital in February 2022 to form NHS England's Transformation Directorate¹) to develop a solution that would provide those national organisations responsible for co-ordinating the response with secure, reliable and timely data – in a way that protected the privacy of citizens – so that they could make informed, effective decisions.

This needed data from across the NHS, social care and partner organisations. In March 2020, a joint team from these organisations and the Department of Health and Social Care (DHSC) came together and worked with the secretary of state for health and social care to issue control of patient information (COPI) notices, which required that data was shared for purposes of Covid.² The health secretary first issued COPI notices in March 2020 and they were renewed several times until June 2022.

Once the COPI notices were in place, the team built the NHS Covid-19 Data Store – a single repository of Covid datasets needed to inform an effective response – on Microsoft Azure. This data is collected from a variety of sources, including NHS 111 online and call centre data, hospital admissions data, testing data and Covid-related deaths data.

NHS England told us that none of the information in the Covid-19 Data Store can be used to identify any individual: identifiable information follows a strict de-identification process where it is anonymised, pseudonymised in line with guidance from the Information Commissioner's Office (ICO) or only made available in aggregate.

To facilitate access to the Covid-19 Data Store, a single 'front door' process – one interface – was set up to manage requests from those applying to access data or wanting to add data to it. Anyone requesting access to data had to be able to demonstrate an involvement in the Covid response and show:

- the purpose – it had to be for Covid purposes
- the type and amount of data requested, with justification – for example, explanations were necessary if anyone requested patient record level data
- transparency – any data requested required a demonstration over how it would be transparently processed
- the legal basis – the legal basis that supported the applicant to process the data.

* The single organisation responsible for managing the NHS and supporting better leadership and care.

** A joint unit of NHS England and the Department of Health and Social Care tasked with driving digital transformation across the NHS.

Although NHS England publishes details of the data involved and its ground for processing data, the National Data Guardian, a statutory body that advises and challenges the health system on its use of data, expressed disappointment at the time taken for greater transparency – specifically, a data dissemination register around how data was used and shared.^{*3}

To provide a “single”⁴ or “shared version of the truth”⁵ about the rapidly evolving Covid situation, data processed via the Covid-19 Data Store needed to be analysed to make it meaningful – which was the function of the NHS National Data Platform. (Covid-19 National Data Solution is sometimes used in this paper to refer to both projects.) Using Palantir’s data platform software (Foundry), dashboards were created to give a live view of the metrics needed to track and understand the spread of the virus, and the health care system capacity available to deal with it. This information supported decision makers, from senior national politicians down to local health officials.

Initially, three dashboards were developed. There was a strategic decision makers dashboard for senior national and regional officials; an operational dashboard for local organisations to make decisions for their local area; and a public dashboard to provide the public with timely updates at national and local levels to ensure good understanding of the progress of the pandemic. These were followed by a Covid-19 early warning system, which provided forecasts to anticipate where the virus might spread next and how that would affect health and social care services; Supply Chain 360, to manage the supply of critical equipment such as personal protective equipment (PPE), ventilators and oxygen supplies; and vaccination capability, to manage the Covid vaccination programme.

According to NHS England, the National Data Platform was instrumental in giving decision makers access to accurate real-time information to make informed, effective decisions. Using the dashboards developed in the Data Platform, decision makers could:

- understand how the virus was spreading and identify risks to particularly vulnerable populations
- proactively increase health and care resources in emerging hotspots
- ensure critical equipment was supplied to the facilities with the greatest need
- divert patients/service users to the facilities that were best able to care for them based on demand, resources and staffing capacity
- ensure fair and equitable uptake of the Covid vaccination.

* The register is available at www.england.nhs.uk/publication/data-dissemination-register

The government and the NHS enlisted the help of several private companies:

- **Microsoft** supported technical teams to build the Covid-19 Data Store on Microsoft's cloud platform, Azure.
- **Palantir** was contracted to provide NHS England and NHS Improvement with Data Platform software.
- **Amazon Web Services (AWS)** provided the cloud infrastructure and technologies that underpin the Data Platform.
- **Faculty AI** helped develop some of the forecasts within the Data Platform, such as the early warning system.
- **NHS Arden and GEM CSU** (on NHS England and NHS Improvement's behalf) operated the single front door for data access requests.

The NHS Covid-19 Data Store website also lists support from McKinsey and Deloitte to "support and help improve the Innovative Data Analytics capacity and capability".⁶ At its peak, NHS England estimates that around 400 people were working to support the Covid-19 Data and Analytics Cell in NHS England, in the joint team created between NHS England, NHSX, NHS Digital and other arm's length bodies, and with consultants and contractors.

The Covid-19 Data Store and Data Platform were cited as a case study in *Data Saves Lives: Reshaping health and social care with data*, the DHSC's data strategy, in June 2022.⁷

One thing participants were proud of

As an icebreaker question, we asked roundtable participants to name one thing they were proud of in their work on the programme. Their answers were:

- The speed and scale of the response to very difficult problems throughout the first six months of the pandemic.
- During this public health emergency, we were able to show our ability to act with agility and resilience by flexing our information governance frameworks, enabling the rapid sharing of information while maintaining proportionate safeguards which deliver on the NHS's constitutional values.
- The way that we worked as a multidisciplinary team across the NHS with partners, and were able to stand up both the data sharing but also some of the functionality we then subsequently delivered across the system – a great team effort.
- I've been in the NHS several years and for the first time ever, we saw attitudes to data, and access to and sharing of data, change at a really rapid pace. There were no barriers: data was seen as accessible to everyone, and not dependent on which NHS organisation you were part of, or the fact that you were an analyst; people who aren't analysts were accessing data, which was really fantastic to see.

-
- How the team continued to work with the trajectory of change and the change in relationships with data, continuing to accelerate the way that we can use data across the NHS to help drive improvements in health care. Also, how the team have not let the pandemic be a waste, learning from it and evolving continuously.
 - Never before had I imagined you would have such a large community of data and data analytics professionals so excited about data and analysis. We're on a very solid footing to make sure that data and analytics begin to inform decision making and we move away from gut-based policy making to more informed decision making. I'm very proud of all the energy and excitement in the system and want to see more of this in the future because we don't always need a pandemic to make us think about the future – we can do this without a pandemic.
 - The nature of the team that came together across organisations and across disciplines was incredible; what was achieved so fast was beyond anything I've ever witnessed before. Just being able to see the impact that was made by data – so tangibly, both operationally and clinically – was just a sight to behold.

Key themes from the discussion

- The clear purpose and urgency that the pandemic brought helped teams to deliver, but the urgency also brought downsides, including staff stress. Routines are now moving towards business as usual and new ways of working (distributed leadership, remote working and collaboration) are being embedded.
- High-level support – from figures including the chief scientific adviser and national statistician, as well as senior figures within the health system – helped empower the team, although business as usual will have to contend with these figures stepping back.
- This support also helped galvanise leadership across different organisations in the health system – but there were challenges in co-ordinating different organisations, duplication of work and overcoming entrenched cultural reluctance to share data between different bodies (legal barriers were much less significant in sharing data).
- Data was brought to the centre of decision making like never before, and its operational use helped to break down silos between data and analysis, and the rest of the system. There was a shift in leaders recognising the need to be data literate (although there is still some work to do). It was also vital that domain knowledge of the NHS supplemented this data knowledge.
- The creation of the Covid-19 Data Store and Data Platform highlighted the importance of preparation: developing an ontology for thinking about the constituent parts of the health system and the data within it early on helped greatly, while a lack of existing guidance and processes for data sharing caused some delays.

-
- The Data Store team tried to engage the public throughout (including through citizens' juries), although communications could sometimes be challenging given the velocity and veracity of stories about the scheme. Government in general needs to do more to engage the public in the use of their data, as a loss of trust – through instances of data sharing the public would not support – would set back progress. NHS England is planning to do more public engagement in future.

Clarity of purpose

Participants felt that **having a clear vision and common goal** that senior leadership set enabled those involved in the Covid-19 Data Store and Data Platform to successfully work together to deliver what was needed. This vision was building a single source of truth and bringing together usable data from multiple sources across the health and care system into a single, secure location for high-quality and relevant information (a "unified front door"). This provided decision makers with a real-time view of the situation (cases, testing, vaccinations), allowing them to make quick and responsive decisions. This was possible thanks to policy work between different parts of the NHS, the DHSC and others including the Office for National Statistics (ONS) and public health professionals, to ensure health professionals could access data (initially through COPI notices).

One participant thought "the biggest thing" was that the Data Store was set up for the sole **purpose of sharing the data**, rather than protecting it (although data protection was built into the design from the start).

The Data Store enabled the NHS to **find the data needed to solve whatever "hard problem" it was facing that day**, from understanding how many people had Covid across the country, to how many ventilators and how much oxygen were available in different hospitals. It meant that the same source of truth across multiple organisations and places underpinned the problem-solving process.

Urgency

Speed was of the essence and drove innovation and improvement: as one participant put it, "we moved on years in months". People within the health system needed to work through 'use cases' – setting out why they wanted the data and how they would use it – with some able to get the data they needed in two weeks, rather than months. (Others waited longer as they were asked to be clearer about their use cases – there was still tiered access, rather than an 'open the gates and everyone has access' approach.)

Something participants found liberating – if also scary – was the need to **do things quickly and work out later if they addressed the problem statement**: "we were throwing the tracks down as we went and you kind of repented afterwards"; there was a sense of "move fast and break things"; "we could say, here's what we think is necessary, let's go with it, and if we need to iterate it we can, so long as we've got a record of

what we've done". Essentially, there was a tolerance of mistakes as part of a process of testing and learning. Sometimes, this meant a lot of effort went into setting up data and knowledge around one subject ("Did we have enough ventilators?") before teams suddenly had to drop it and rapidly move on to the next primary concern.

But there was **no time to think, or for horizon scanning**; the urgent need meant day-to-day decision making was prioritised. This had drawbacks – one participant said that a greater degree of thought leadership may have been useful given other countries wanted to build on the Data Store and Data Platform experience. But these considerations were less urgent. One participant told how a clinical leader rang up on a Friday afternoon, saying: "We need to do the following things over the weekend otherwise people will die" – that "brought home a new dimension to my job I've never had before. I didn't do much horizon scanning that particular weekend, I recall."

'Piggybacking' on existing routines like **daily situation reports ('sitreps')** helped **mitigate some of the downsides of speed**. While not horizon scanning, these were set up in such a way that the Data Store and Data Platform team could understand what they needed before the next cycle of data to help answer particular questions, think through how else the data they were collecting could be useful, and even identify data gaps (for example, around mental health).

Staff

The team went through "significant" stress and there is a need to move away from emergency operations to a more sustainable model so the NHS "[doesn't] burn people out like we did". This includes adapting working patterns – people worked in daily sprints – but also resetting expectations. Some NHS England teams suddenly had a responsibility to keep data flowing over a weekend and respond immediately to requests for information; some people requesting data retain that expectation even though such speed is no longer necessary. One participant said a colleague on call over a weekend had missed four calls at 2am and felt they should respond to a voicemail at 5am, even though the response should now be to wait until Monday. Those asking for the data have keenly felt the loss of immediate service, while staff have also "found it hard to step back because the responsibility has been on their shoulders for too long" and because of their passion for getting things done.

There have been **several changes to ways of working**. Distributed leadership meant a move away from a hierarchical paradigm of all decisions having to come from the top (meaning many decisions would never get made). Instead, Ming Tang (NHS chief data analytics officer) and her team were praised for empowering people at all levels. Junior staff now want to work differently; working more with the private sector and the third sector (and feeling empowered to have the necessary conversations with these external partners), and quickly working up proofs of concept rather than having endless stages of planning. Remote working "levelled up the playing field in terms of who could play on the pitch" (although early on, many people were still going in to the office). Covid

opened up conversations where senior people and colleagues on the front line were meeting to discuss key issues – participants felt it important that this is not lost. A more federated, distributed approach means the NHS is “trying to make the best use of everyone wherever they sit”, and it now “expect[s] people to work wherever they need to be to be closer to the work”. Staffing models have changed as the number of consultants involved has been reduced (with consequent upskilling of NHS staff and significant recruitment), as have working patterns – for example, moving to shifts given the seven-day weeks and longer hours that staff had been working.

Domain knowledge and data literacy

Multidisciplinary teams bringing together different expertise were critical.

Participants said hours were wasted as talented data scientists – who wanted to help and whose expertise was critical – worked on data without understanding the context of how it would be used, given NHS data is very ‘dirty’ and contextual (this also created some fear in giving it to them). Co-production, bringing the business process and the analysis together, worked: daily sit-downs with people running the emergency response meant everyone could understand what data models were saying, and allowed triangulation of the data with the experience on the ground – since data does not always give the right answer. Another example was of directors or policy professionals needing to win arguments with political audiences to unblock problems their teams faced: they do not need to know all the detail their teams know, but need enough from their analysts to boil down key points without losing critical nuance. Agreement on sharing data upfront would have helped further, and the NHS is now trying to place analysts closer to where the data is being used to improve understanding.

The **role and status of analysis have shifted** during the pandemic. The vital importance of analysts’ work on the Covid-19 Data Store meant they took a more prominent role, at the forefront of organisational change and operational problem solving, rather than being seen as a back-office supporting function. ‘Data’ was not seen as siloed. As one participant put it: “There is no such thing as a ‘data person’, there’s no such thing as a ‘data strategy’... everything was happening cohesively, which was really beautiful.” At the same time, directors across the health service now talk about data in a way they did not before, and people want to work together around data. “There has been a real change in the way that people see the value of data and how that can help them do their day job.” This relies on a greater level of data literacy from senior leaders. One participant said that, for too long, “it’s been acceptable for very senior people to say, ‘I’m no good with numbers’”; the pandemic has helped shift that culture, although there is still further to go. Some attendees felt that all the data, dashboards and analysis in the world would fail without senior leaders being data literate, which is what supports more effective decision making.

Jessica McEvoy, now a principal consultant at software consultancy Scott Logic and formerly deputy director (Platforms and Services and head of GDS International) at the Government Digital Service (GDS), said: “There’s real potential to build on these lessons from the pandemic and foster data-imbued organisational cultures. For that to happen, we need data-literate leaders who spurn the siren song of off-the-shelf technical solutions. What’s important instead is to get the foundational stuff right – having a clear organisational strategy and then investing in the right people, technology and skills to enable you to harness data to deliver that strategy. Departments need more than just a data strategy; they also need to know how data can empower their missions and how to use data to deliver their overall departmental strategy – which means being clear on their objectives and understanding what questions they need data to answer.”

Political leadership and high-level buy-in

The Covid-19 Data Store and Data Platform had the **backing of senior figures in government** – including Sir Patrick Vallance (chief scientific adviser) and Sir Ian Diamond (national statistician). This helped bring about a “congruence of leadership at all levels” never seen before, and helped empower figures like the NHS chief data and analytics officer (the “tip of the spear”) to galvanise analysts and data professionals – “a game-changer”. A “huge amount” of senior attention was essential in the early stages to facilitate data sharing decisions; the ongoing model must account for not having that, now the original crisis has subsided and those figures step back.

Nonetheless, the Data Store and Data Platform team sometimes felt **stuck in the middle between different organisations** – for example, with the DHSC and NHS England wanting to maintain sovereignty over Number 10 and the Cabinet Office. Different organisations were asking for different data and wanting to use it in different ways, which was stressful and wasted time. Eventually, the Data Store team built an application programming interface (API), allowing different requests to be made of the same underlying data and some automation.

The prevailing cultural reticence to share data also created problems, which meant different parts of the health system took time to work through what they were comfortable sharing, delaying valuable insights. Additionally, a ‘knowledge is power’ culture meant that government departments, reliant on successful spending review bids to the Treasury, were reluctant to share data given concerns about how the data would be interpreted and used (what would the consequences be if activity did not match spending review numbers?). That said, participants felt that organisations were learning to share data more freely, although questions remained about which organisations would be accountable for decisions based on shared data.

Relationships and working across organisations

One participant compared the number of organisations involved to “primary school football” – initially **“everyone was chasing the ball”**, with no sense of tactics or co-ordination. There was duplication of work between different bodies – the NHS, the Joint Biosecurity Centre and the former Public Health England – which one participant described as “madness”. It took time to resolve these issues, as teams had different ways of doing things and multiple bodies believed that they were in charge of the same thing.

Senior figures like Patrick Vallance saying the Data Store team was “in charge” helped – they then held cross-system meetings to gather requirements from different groups and make the data available. Organisational boundaries disappeared, as jostling, self-preservation and the guarding of turf (“these are *my* patients and *I’m* going to deliver care for them”) gave way to collaboration around the patient. A more federated data platform democratised access for national teams and local ones, fostering collaboration between analysts as they shared code and insights and partnered with other sectors. Participants felt that the NHS moved from looking inwards with a strong sense of ‘not invented here’, to seeking out other people who do things better and becoming a learning system. More openness and greater awareness about what everyone is working on, better sharing of best practice and greater clarity as to who is responsible would help further.

These dynamics shaped the technical build of the Covid-19 Data Store. People like to be in charge of an outcome or output, which makes trying to build something that does many things difficult: it will never quite do exactly what individual business units want it to do, and you cannot anticipate exactly how they will use it or the value they will generate. The team working on the Data Store therefore made their technical build as small as possible – exposing things in their rawest form, safely and with the right governance – which participants described as a big shift and unusual in government. The Covid-19 data response team also used a cloud approach (that is, making flexible use of existing on-demand computing services), which they say simplified the infrastructure, reduced capital expenditure and allowed the team to prove value quickly. Automation was used where possible to streamline data flows, reduce data latency and ensure data could flow to those who needed it effectively and efficiently.

There was a helpful general acceptance that the private, public and third sectors needed to work together, even if more could have been done with partnerships with the private sector, third sector and other large initiatives (such as the ONS, HDR UK, DECOVID and the Emergent Alliance*). These partnerships provided the ability to scale up rapidly and deliver quickly. Relationships improved as a clear model for collaboration developed, working together delivered successes and it was clear who was in charge.

* HDR UK – Health Data Research UK – is the national institute for health data science; DECOVID is a database of health information; and the Emergent Alliance is a not-for-profit initiative of private companies and others, which created a safe environment for data sharing.

The Covid-19 Data Solution was able to provide a mechanism to share and support other studies and analysis, like REACT,* so they could operate more efficiently. The “everyone-pitching-in, as-fast-as-possible, attitude” meant some private contractors led more than they might have expected; the NHS and private partners have moved away from that with the NHS now more clearly taking the lead.

The primacy of data

The fact that **data brought people together** to make decisions was important and “felt very real”. People being able to access one shared version of the truth, directly in real time, was positive: there was not any rehashing of the data or having to retrieve it from a slide. People felt comfortable working with data and wanted to do so because it was so timely, moving from weekly situation reports with data six or eight weeks old before the pandemic, to daily situation reports. The vaccination programme had an hourly feed, which also helped ensure the vaccination process was running smoothly.

The Covid-19 Data Solution was **able to iterate rapidly**: collecting data every day meant the team could make changes if necessary. This had a downside – it can be unhelpful if you are trying to build a consistent time series – but it helped the team understand the data they were getting and how it was being used on the ground; “we don’t usually play back the data we collect to the service in this way”. The daily cadence was important for prompting questions like “what else do we need to do with the information?” and “what’s the next decision we need to make?”.

Information governance (ensuring data is used appropriately) and the structures and routines around it were at the core of everything. A daily meeting brought together the most important people in the NHS and information governance to work out what needed to be done. This meant data could be shared correctly and quickly to solve urgent problems.

Advance preparation and thinking about the future

Getting the ontology right up front was critical and the degree to which the Covid-19 Data Solution was built from an early stage with future data use in mind was “impressive”. As disparate data sources were coming in from hospital wards at the start of the pandemic, data pipelines were sketched out on a white board at NHS England at one of its offices, Skipton House. This is when the ontology – a canonical, conceptual framework for thinking about different entities and objects (like ‘hospital’) and data sources – was built. Having it in place from the start underpinned everything else: different data products across different teams for different purposes – the strategic decision makers’ dashboard, ventilator and PPE allocation and the vaccination programme, for example – were based on the same ontology, meaning the team could quickly solve new problems as they emerged and did not have to start from scratch.

* REACT – REal-time Assessment of Community Transmission – studies, which Imperial College London and Ipsos UK run, aim to understand the level of Covid infection across the country.

For instance, staff could see what was available at individual hospital level without having to know the details of other datasets further upstream. Scalability and evolution were built in: additional datasets that later became available could be folded in to the ontology and shared through the same processes and patterns.

Nonetheless, **there were challenges with the data**, including about the structures of the health system (which one participant described as the “metadata around the truth of the NHS”). NHS organisations are not static, with trusts changing and merging, so data sometimes stopped coming through as the data pipeline broke when organisations changed. There was also “quite a lot of pain” in dealing with competing geographies, with different types of organisation having overlapping boundaries. At one point, the team were asked to provide data at a local resilience forum level – “we might be able to, if we knew what a local resilience forum was”.

There were **delays where things had not been thought through in advance**. These included debates about de-identifying and pseudonymising the data, what could be done with it and how to adhere to people opting out of their data being used by the health service; and data sharing protocols between different organisations (particularly between national and local government), since it felt like there were no established methods (for example, agreed API standards)* for sharing data effectively and securely across government.

Finding and accessing data

Signposting – helping to direct people to go where they needed to access the data – could have been better. Participants felt they could have taken a step back and asked: “What are we trying to do here and how can we help people who are trying to use the data?” There wasn’t a routine understanding of who should see what, so the team set the Data Platform up to allow different levels of access (users being able to see things at national level, regional level or just their own data). The team also significantly shortened the time to access data compared with ‘peacetime’ (although they acknowledged the experience probably felt clunky for some people asking for data). The process has continued to improve, moving from an ad hoc approach to access, run by a daily Skype meeting, to standard operating procedures. It is now easier to understand what data there is, what features there are and how to apply for access, which is more efficient and transparent.

COPI (control of patient information) notices

COPI notices did not solve everything. There were people who would still not hand over data, and initial legal uncertainty was an inhibiting factor. Some people feared they were doing the wrong thing, even though legal bases already existed for much of what they wanted to do. There is a need to break down barriers, so people understand what the law allows.

* The GDS published some API standards in 2018, which the Central Digital and Data Office (CDDO) now oversees. See CDDO, ‘API technical and data standards’, GOV.UK, 2018, retrieved 1 February 2023, www.gov.uk/guidance/gds-api-technical-and-data-standards.

But **COPI allowed the rapid deployment** of the Covid-19 Data Store while ensuring the privacy of the population was protected. Data sharing agreements were put in place quickly rather than taking months. Streamlining processes – such as data sharing agreements, working transparently and collaboratively, flexing the information governance framework and calls (often daily) with different organisations – helped move conversations forward. COPI was like “a badge of honour”, which helped people understand they were covered and opened up conversations with GPs and trusts; “it’s a security blanket, if you like, that sharing was safe”.

Public engagement

Due to the pace at which the programme was working, NHS England told us that a full public consultation was not feasible within the timescales. But NHS England and the government recognised the need to do more to educate the general public about how and why their data is used, and understand their expectations around transparency and how they would like to be involved in the future.

The team did **a lot of engagement with the public through national stakeholder partnership groups** to make sure people understood early on what the Covid-19 Data Store was trying to do and get their input. They also created a public-facing dashboard, which was used for all Number 10 briefings and widely accessed. They felt it was better for the NHS, rather than politicians, to lead the engagement, although it was “tricky” in places – they were always worried they might get shut down.

Between March and May 2021, NHS England and NHS Improvement and NHSX **held three citizens’ juries**, supported by the NIHR Applied Research Collaborative and the National Data Guardian. The juries heard evidence from a variety of expert witnesses and deliberated together to answer a set of jury questions about pandemic data sharing initiatives, including the Covid-19 Data Store and Data Platform. Overall, the juries supported the decisions made to introduce the initiatives during the pandemic, reaching 87% who were broadly or very much supportive of the data initiatives as part of tackling Covid. But there were concerns that there was a lack of transparency about data sharing initiatives.

NHS England told us that it is **committed to being open and transparent** with the public about how their data is used and to giving them choices in accordance with their rights. It is planning to procure a ‘Federated Data Platform’, which includes plans to: publish patient-facing information about who is accessing data and for what purpose (developed in collaboration with medConfidential, UseMyData and the Patients Association); ensure that patients are informed of their ability to opt out; and actively engage the public in decision making processes.

Those working on projects like the Covid-19 Data Store **need government to help them maintain public trust**. As soon as government starts talking about use cases that the public will not accept – for example, joining up health data with data from

the Home Office – “we will lose”. What people want their data to be used for, and who they trust with it across government, is a “fine line”. Participants felt that many people currently do not trust the government to do the right thing with their data. The sharing of personal data should be aligned with society’s values (the Nolan principles for standards in public life were given as an example) and people’s expectations (there should be a blueprint for incorporating the public) to gain public consent and ensure data use is trustworthy.

Communication

The administrative burden on the communications team should not be underestimated, and it was difficult to balance that with continuing to create lots of products. The fact that big data programmes have failed in the past means the shorthand easily becomes “this is the National Programme for IT, or care.data, again”, without people knowing the detail.* The team found some of the coverage frustrating and not always founded on facts. While the team worked hard to manage the noise and provide more nuance around what the Covid-19 Data Store was trying to do, some earlier mythbusting activity might have helped further. Their experience to date means their communications and engagement plan is now clearer, particularly around the benefits and what they are (and are not) planning to do, as they respond to renewed interest in the Data Store as it is put on a more substantive and sustainable footing.

NHS England has since told us that **public distrust over the use of data is a recognised strategic risk** for the Federated Data Platform and that key lessons have been learnt from previous data programmes. The communications team have identified the need for:

- proactive public involvement and engagement
- open communications and transparency throughout the process
- clarity on accountability, governance and safeguards in order to mitigate the risks
- educating the public, providers, clinicians and management on the value the programme plans to deliver.

The DHSC, NHS England and NHS Digital are also planning a single public confidence campaign, bringing messaging together from multiple data programmes, “underpinned by patient and public involvement and engagement, co-creating solutions with the public”.

* The National Programme for IT (NPFIT) was an attempt to improve information storage and sharing across the NHS, including a patient record system, which was scrapped in 2011 after years of delays, technical and contractual difficulties, and increasing cost. Care.data intended to bring individual GP records together so they could be used for research and planning but the programme was halted in 2014 after controversy about the lack of information given to patients. This was one reason why the Major Projects Authority rated the project ‘red’ (successful delivery appears unachievable).

The wider health system

The lack of established and practised data sharing between the devolved nations was a real block to building a UK-wide picture. The team did not pursue some analysis and other possibilities because trying to transfer data between nations “seemed to be just one hurdle too many” given the time pressure. When building data pipelines, the NHS England data teams were not allowed to do this beyond the English border. They made their code available to other teams and encouraged them to use it, but there was no mechanism or body to convene a national conversation.

Social care data was a problem, again partly because there was no national body to negotiate on behalf of the different parts of the system. During the vaccination rollout, for example, the Covid-19 data response team wanted to understand what was happening in care homes but could not.

There was also a lack of information on mental health. The focus on capacity for acute care was understandable, but mental health was not prioritised early enough and the impending impact on it recognised sooner. ‘Parity of esteem’, which means mental health and physical health are valued equally, meant mental health providers were being asked the same questions as others in the health system, but mental health experts argued there was a need for different questions as well.

The future of the Covid-19 Data Store

The Covid-19 Data Store was a temporary project to support the Covid response. The data was due to be destroyed, but as the data sharing agreements have moved to a non-COPI basis, **it has been possible to sustain the data** in the National Data Platform. Having seen that “these datasets in proximity to each other does spark ideas”, there is a strong desire to increase the use of data to (for example) inform, forecast and plan around demands on the health care system. Learning from the Covid-19 Data Store has “opened the doors to all sorts of potential future use cases that are very high impact” and will be put into practice supporting the Data Platform and future data initiatives.

According to NHS England, it is **planning to invest in digitising data** – for example, the DHSC’s data strategy *Data Saves Lives*⁸ commits to building on the capability developed over the past couple of years (using the Covid-19 Data Platform), including modernising its data architecture through the new Federated Data Platform. This aims to bring together operational data (currently stored in separate systems), including the number of beds in a hospital, the size of waiting lists for elective care services and the availability of medical supplies, to support staff and connect and share information between every hospital trust and integrated care system.

Key lessons from participants

Participants drew out several key lessons for government based on the Covid-19 Data Store and Data Platform experience. These included:

- establishing methods for data sharing, and shared processes, which would provide clarity around the future of the data landscape and ensure collaborations can operate efficiently and confidently
- introducing a legislative 'duty to share' during emergencies that could help make appropriate data sharing the norm rather than the exception
- having greater openness within government (including having regular meetings, sharing best practice and being clear about organisational responsibilities, to help see what people are working on and avoid duplication) and outside (to foster innovation and collaboration)
- maintaining public trust through aligning the sharing of personal data with society's values (for example, the Nolan principles) and people's expectations – the latter will require a blueprint from government for incorporating the public in setting the purpose and agenda for data sharing projects, and maintaining patient, service user and community engagement throughout the life of the projects
- establishing a communication blueprint across the health system for the effective and efficient devolution of policies, guidance and communications around data
- tidying up legislation (such as the Digital Economy Act 2017 and Data Protection Act 2018), since many difficulties around legal bases and information governance come from the way different Acts come together and sometimes contradict one another
- having common geographies (that is, consistent and clear boundaries) for organisations in the health system, or at least raising understanding of the different geographies and types of local organisation and their needs, to make it easier to collate and provide data for them
- supporting international collaboration to establish the mechanisms for full data sharing across countries in the event of a future pandemic (either a single, global international registry or a federated network underpinned with clear mapping and standardisation)
- setting an even bolder vision around the citizen data scientist (for example, during the pandemic, many citizens have reported Covid symptoms using apps).

Gavin Freeguard is a freelance consultant and former programme director at the Institute for Government.

Paul Shepley is a data scientist at the Institute for Government.

References

- 1 Gould M, 'NHSX moves on', NHS England blog, 4 February 2022, retrieved 1 February 2023, <https://transform.england.nhs.uk/blogs/nhsx-moves-on>
- 2 NHS Digital, 'Control of patient information (COPI) notice', (no date), retrieved 1 February 2023, <https://digital.nhs.uk/coronavirus/coronavirus-covid-19-response-information-governance-hub/control-of-patient-information-copi-notice>
- 3 National Data Guardian, *The National Data Guardian for Health and Social Care: Annual report 2020–2021*, 2021, retrieved 1 February 2023, www.gov.uk/government/publications/national-data-guardian-2020-2021-report
- 4 Department of Health and Social Care, *Data Saves Lives: Reshaping health and social care with data*, GOV.UK, 2022, retrieved 1 February 2023, 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
- 5 NHS Digital, 'NHS National Data Platform', (no date), retrieved 1 February 2023, <https://digital.nhs.uk/coronavirus/vaccinations/training-and-onboarding/reporting/nhs-national-data-platform>
- 6 NHS England, 'NHS COVID-19 Data Store', (no date), retrieved 1 February 2023, www.england.nhs.uk/contact-us/privacy-notice/how-we-use-your-information/covid-19-response/nhs-covid-19-data-store
- 7 Department of Health and Social Care, *Data Saves Lives: Reshaping health and social care with data*, GOV.UK, 2022, retrieved 1 February 2023, 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
- 8 *Ibid.*

The Institute for Government is the leading think tank working to make government more effective.

We provide rigorous research and analysis, topical commentary and public events to explore the key challenges facing government.

We offer a space for discussion and fresh thinking, to help senior politicians and civil servants think differently and bring about change.

Copies of this IfG Insight are available alongside our other research work at:

 instituteforgovernment.org.uk

 enquiries@instituteforgovernment.org.uk

 +44 (0) 20 7747 0400  +44 (0) 20 7766 0700

 [@instituteforgov](https://twitter.com/instituteforgov)

**Institute for Government, 2 Carlton Gardens
London SW1Y 5AA, United Kingdom**

February 2023

© Institute for Government 2023

The Institute for Government is a registered charity in England and Wales (No.1123926) with cross-party governance. Our main funder is the Gatsby Charitable Foundation, one of the Sainsbury Family Charitable Trusts.