NDG National Data Guardian
for health and social care

The National Data Guardian's written submission to the Science and Technology Committee's call for evidence on 'The right to privacy; digital data'

7 February 2022

Contents

Introduction

This is the National Data Guardian's (NDG's) written submission to the Science and Technology Committees' call for evidence on 'The right to privacy; digital data'.

The role of the National Data Guardian

The National Data Guardian for health and adult social care in England (NDG) is appointed by the Secretary of State for Health and Social Care to serve as an independent champion for patients and the public when it comes to matters of their confidential health and social care information. The NDG role was introduced in 2014 to build public trust in data use by advising, encouraging and challenging the government and those who work within the health and social care system to ensure that people's confidential information is being kept safe and secure, and only being shared when appropriate to achieve better outcomes for patients and the public.

The present NDG is Dr Nicola Byrne, a consultant psychiatrist in adult mental health, who has held the role since April 2021. Previously, the role was held by Dame Fiona Caldicott.

Which elements of the inquiry this response addresses

The NDG would like to commend the committee for initiating an inquiry into this vast and complex topic. This response does not address all areas set out for exploration in the inquiry, only those that fall under the NDG's remit. This inquiry asks about sharing data across a wide range of different organisations such as 'government departments, other public bodies, research institutions and commercial organisations'. Other questions ask about sharing within discrete contexts such as 'health and care contexts'. Given the NDG's remit, this response only addresses the sharing of health and adult social care information. Should the committee wish to explore any of these topics further, the NDG would be willing to provide oral evidence to the committee.

Terms relevant to this submission

For clarity of language, the key concepts and programmes considered in this response are set out below:

Personal data

Personal data is defined in the UK GDPR as:

"'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

This means personal data has to be information that relates to an individual. That individual must be identified or identifiable either directly or indirectly from one or more identifiers or from factors specific to the individual.

Confidential Patient Information (CPI)

Patients using health and social care services are entitled to expect that their personal information will remain confidential. They must feel able to discuss sensitive matters with healthcare professionals without fear that the information may be improperly disclosed. These services cannot work effectively without the trust that depends on confidentiality.

Patients have a reasonable expectation that the information they share with their care teams will not be used for purposes beyond their own, individual care without their consent (save for certain permitted secondary uses to support the health and social care system). The duty of confidentiality requires that where confidential patient information is used for purposes other than an individual's own care and treatment, there must be either explicit consent, an exemption provided by law or an overriding public interest. If there is an intention to access confidential patient information without consent in England and Wales, organisations should apply to the Confidentiality Advisory Group (CAG) for section 251 support. This is a shorthand term and refers to section 251 of the National Health Service Act 2006 and its current Regulations, the Health Service (Control of Patient Information) Regulations 2002. The NHS Act 2006 and the Regulations enable the common law duty of confidentiality to be

temporarily lifted so that confidential patient information can be transferred to an applicant without the discloser being in breach of the common law duty of confidentiality.

Anonymous information

Anonymous information is information that does not relate to an identified or identifiable individual. Neither data protection law nor the common law duty of confidentiality apply to data truly rendered anonymous in such a way that the data subject is no longer identifiable. This enables anonymous information to be used more widely in the health and care system for secondary uses, such as research and planning.

The Caldicott Principles

The Caldicott Principles are eight well-established, good practice guidelines that apply to the use of confidential information within health and social care organisations and when such information is shared with other organisations and between individuals, both for individual care and for other purposes. They are intended to apply to all data collected for the provision of health and social care services where patients and service users can be identified and would expect that it will be kept private. This may include for instance, details about symptoms, diagnosis, treatment, names and addresses. In some instances, the principles should also be applied to the processing of staff information. They are primarily intended to guide organisations and their staff, but patients, service users and/or their representatives should be included as active partners in the use of confidential information.

Legal frameworks for health and care data use

Data collected in different contexts will be subject to different legal frameworks, which require different procedures and safeguards. Whilst the UK GDPR and the Data Protection Act 2018 provide a national approach to data protection, the duty of confidentiality developed through the common law, which applies to confidential patient information, only applies to health and care data.

Thus, health data requires different treatment and safeguards based on this distinct legal regime. As this legal regime only applies to health and

care data collected in the context of a confidential relationship, its operation and legal bases for lifting the obligation of confidentiality are not always well understood in non-health and care contexts.

Given the unique status of health and care data, any proposals considered by the inquiry to share confidential patient information across sectors should be approached very differently to the data sharing within the health and social care system also considered by this inquiry.

Key considerations for health and care data use

The relational context of health and care data collection and the critical importance of trust

It is important to note that in addition to the distinct legal regime protecting confidential health and care data, by its very nature it is uniquely sensitive. It is shared by patients and service users with their health and care professionals on the basis of trust, which is founded on an expectation of confidentiality and a respect for privacy. Because of this, people expect significant safeguards and strict limits to its use.

They need to be able to trust that those who collect this data about them will respect it. If people do not feel able to trust that their information will be kept private or used appropriately, this may cause harm: it could affect their willingness to seek care; or when seeking care, their willingness to provide full and accurate information – which would be to the detriment of the safety and effectiveness their treatment. Such a scenario would also harm system planning, research, and innovation, which rely on the availability of full, accurate and representative data.

There is a wealth of empirical evidence which demonstrates that whilst people have a high level of trust in the NHS to use their data properly and keep it secure, they do not have the same level of trust in other organisations, such as government departments, research organisations and commercial organisations. For evidence of this, please see: Sharing anonymised patient-level data where there is a mixed public and private benefit - a new report - Health Research Authority (hra.nhs.uk), Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data - GOV.UK (www.gov.uk).

Trust is an increasingly prominent theme in discourse around data. However, it is not enough for any organisation to just *say* that it recognises the importance of public trust: it must *demonstrate* its trustworthiness in order to earn it.

A recent case in point is the negative public response to the General Practice Data for Planning and Research (GPDPR) programme. GPDPR is an NHS Digital programme that is making improvements to how data is collected from general practice, with the introduction of a new framework for data extraction called the GPDPR collection.

Its aim, which the NDG supports, is to collect the data held in the GP medical records of patients and ensure that it is used every day to support health and care planning and research in England, helping to find better treatments and improve patient outcomes for everyone.

The data collection was due to start in summer 2021 but was delayed after significant public concern and media criticism that focused on the lack of public and professional engagement about the programme. People had not felt informed or reassured and this created fear, which translated into a rise in the number of people opting out of their data being used for research and planning. The then Parliamentary Under Secretary of State, Jo Churchill MP, announced the delay in a letter to GPs; the letter set out criteria that would need to be met before the programme could proceed (aimed at improving safeguards and engaging with people to raise awareness and build trust).

This demonstrated that whilst the public's data literacy and awareness of the benefits of data use may have increased during the pandemic, its trust in the NHS's ability to keep their data safe and use it appropriately diminished, arguably in tandem with growing public concerns about both the potential risks and commodification of data use at scale.

Any plans by the government to routinely share 'data between and across government, other public bodies, research institutions and commercial organisations' risks damaging fragile public trust even further. Data collected for the purpose of providing care must be given the utmost respect. It must be used appropriately, in circumscribed ways, in accordance with people's expectations, so that the previously described harms do not come to pass.

Acknowledging risk

This call for evidence asks respondents to expand upon the benefits of, and barriers to, sharing data. Whilst this is important, it is equally

important to acknowledge that sharing data, particularly that which is capable of identifying individuals, carries risks. Failing to be honest and transparent about the existence of those risks (and how they will be mitigated) arguably results in the most significant barrier to sharing of all: a suspicion of deceit, and an ensuing lack of trust. Organisations that do not demonstrate a genuine commitment to openness and transparency about data use (including an acknowledgement of risk) are unlikely to be judged by the public as trustworthy.

We have seen what happens when the importance of transparency and engagement is underestimated. Where initiatives to share people's health and care data are not properly explained or understood (by both the public and professionals) this causes suspicion, mistrust and fear – to an extent that can not only derail the initiative (as with the care.data programme), but can also deal a longer-term blow to people's trust in the confidentiality of our health and care services.

Protecting privacy through anonymisation

This inquiry considers two types of data: data that has been rendered anonymous, and data that is capable of identifying individuals.

It is important to note that the legal responsibilities and risks relating to each of these two types of data are different. Data rendered anonymous is not data about individuals and is therefore not protected by either the data protection regime or the common law duty of confidentiality. This is because the risk to privacy where anonymous data is shared is not the same as with data that is capable of identifying individuals. Data that is capable of identifying individuals *is* subject to both the data protection law and the common law duty of confidentiality.

Determining whether data is truly anonymous is complex, and the term can be subject to varying interpretations. Recent consultations held by the Information Commissioner's Office on its proposed anonymisation guidance have begun to provide clarity in this area. Please see: ICO call for views: Anonymisation, pseudonymisation and privacy enhancing technologies guidance | ICO.

The inquiry will need to determine and be clear about where it is considering anonymous data, and where it is considering data capable of identifying individuals, so that the appropriate considerations can be made, and conclusions drawn.

The NDG is supportive of mechanisms that render health data anonymous, where this is possible, in order to protect people's privacy. For more about

the NDG's historical support for the use of anonymous data, please see: Review of data security, consent and opt-outs – GOV.UK (www.gov.uk).

The NDG is also supportive of provisions to allow health and social care bodies to share data that has already been rendered anonymous with other health and social care bodies to support the provision of health and social care services.

Even though sharing anonymous data may not present risks to *privacy*, people perceive other significant risks and harms that may result from it, and only support its use where it is for public benefit. When anonymous data is used for reasons that the public do not consider to be for public benefit, it can damage their trust in the integrity of the system, and its ability to use data appropriately. For evidence supporting this point, please see: Sharing anonymised patient-level data where there is a mixed public and private benefit - a new report - Health Research Authority (hra.nhs.uk).

Where data capable of identifying individuals is shared, public attitudes tend to be more conservative and guarded because people also perceive a risk to their own privacy. For evidence, please see: Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data - GOV.UK (www.gov.uk).

Any policy relating to sharing health and care data should reflect public concerns about both privacy and use.

Data sharing benefits

There are great benefits to sharing information across the health and social care sector:

- enabling the best delivery of care to individual patients
- learning from people's experience of health, illness and treatment, including what works and what doesn't, and to identify both good and poor practice so as to improve health and care for others in future
- operational planning to support the most effective use of our public finances, to deliver a sustainable health and social care system that benefits everyone
- undertaking research and innovation to improve health and disease management through developing new approaches to prevention, treatment and more person-centred care

Data sharing for individual care (direct care)

There are clear benefits to sharing confidential patient information in support of people's individual care. The NDG recognises this in Caldicott Principle 7:

The duty to share information for individual care is as important as the duty to protect patient confidentiality:

Health and social care professionals should have the confidence to share confidential information in the best interests of patients and service users within the framework set out by these principles. They should be supported by the policies of their employers, regulators and professional bodies.

This principle is supported by Section 251B of the Health and Social Care (Safety and Quality) Act 2015 which provides that there is a duty to disclose information 'likely to facilitate the provision to the individual of health services or adult social care.'

The legal basis that permits data sharing for the purpose individual (or direct) care is 'implied consent'. Patients expect their information to be accessed by those treating them, and therefore their consent can be presumed from that expectation. Data sharing for individual care is limited to those within a patient's health and care team, who have a 'legitimate relationship' with that person (and therefore a need to access their information to treat them). This limitation should not be seen as a barrier. It should be recognised as an important limit on what sort of data sharing can rely on implied consent as its legal basis. It is a necessary boundary imposed to maintain patient trust in health professionals.

However, the boundaries of direct care are often not understood in practice, which can cause confusion about whether the purpose for which the information being shared is, in fact, direct care (which has ramifications for the legal basis).

In 2020, the NDG carried out a survey to identify the key barriers that prevent health and social care staff from sharing information appropriately in support of direct care. The findings are published here: NDG report on barriers to information sharing to support direct care - GOV.UK (www.gov.uk). The NDG made four recommendations to reduce the barriers identified:

1. an education and training strategy to encourage information sharing for individual care

2. greater clarity about what falls within individual care and what does not
3. development of an approach to ensure that patients, carers and service users can access important information about their health and care
4. a better understanding of what specific data and information is required by the health and care system to meet the different demands of care provision, research and planning.

The NDG also developed a draft decision-support tool to help health and care professionals determine whether an activity falls under the banner of direct care or not: NDG report on barriers to information sharing to support direct care - GOV.UK (www.gov.uk).

Data sharing for service planning and evaluation

Data collected in the context of providing health and care can also be used for 'secondary purposes', such as the planning and evaluation of NHS services. However, where confidential patient information is used outside of individual care for these secondary purposes, it should be subject to safeguards that reassure the public they can trust the health and care system to use their data safely for reasons other than their own care.

Using anonymous data

Where anonymous data can be used for planning purposes instead of information capable of identifying individuals, this is preferable. The NDG supports the principle of using anonymous data (that is properly anonymised in line with ICO guidance) to plan and evaluate essential NHS services, as this minimises risks to individual privacy.

Health and care data is our shared public asset, and the safe and effective use of this asset can deliver huge benefits for the common good. Viewing anonymous data as a shared asset, rather than something that is exclusively 'owned' by individuals or by the system, could build public trust through engagement and involvement in its use, and in decisions made about it. As a shared public asset, when new ways of using data are proposed, engaging the public as active stakeholders through public representation on oversight and decision-making forums, would be one way of increasing public confidence in the governance of its appropriate, equitable use for public benefit.

Although privacy risks are minimised wherever anonymous data is used, other risks perceived by the public still exist. For example, people have

strong views and legitimate concerns about the ways in which anonymised health and social care data might be used when shared outside of the health and care system.

This highlights the 'ownership' tension that can sometimes arise from public sector data use. Empirical research demonstrates that where anonymous data that was provided by the public for health and care use is then used for secondary purposes, the public believes that those secondary purposes must deliver a public benefit. For an example, see: Sharing anonymised patient-level data where there is a mixed public and private benefit - a new report - Health Research Authority (hra.nhs.uk), Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data - GOV.UK (www.gov.uk). The need for data use to deliver public benefit should not be seen as a barrier: it is a necessary safeguard, and one which the public has repeatedly stated that it requires for the system to demonstrate it is trustworthy.

Data sharing for research

Data collected for health and care can also be used to perform vital research, which can improve health and care for everyone.

The practices for gaining access to data collected during care for research purposes must be robust if they are to earn public trust. If people do not have trust in how the health and care system uses their data for research, they are more likely to choose to exercise their right to opt-out of sharing their data for this reason. A significant rise in the number of people opting out may affect the quality of the data and the research that can be undertaken with it.

Important safeguards when data is used for research

There are a number of important safeguards that govern access to healthcare data for research. These should not be considered as barriers, but rather respected for what they are: mechanisms for earning and maintaining trust in the sharing of health and care data for research.

Oversight groups

Where a programme wishes to use health and care data in its research, the matter should be considered by a Research Ethics Committee (REC). Where the research seeks to data collected during care delivery for the purposes of research, CAG advice and HRA approval should be sought under NHS Act 2006 and the COPI Regulations to enable the common law duty of

confidentiality to be temporarily lifted so that confidential patient information can be transferred to an applicant without the discloser being in breach of the common law duty of confidentiality. CAG is a lay and expert committee which provides advice to the HRA on the use of health data without consent for research under Section 251 of the NHS Act 2006.

Scrutiny and challenge from oversight organisations such as research ethics committees and the CAG should not be seen as barriers to research using data collected during care provision. People want to be able to trust that where health and care data about them is used for research, it is for the public's benefit and that their data will be treated appropriately and securely. This is crucial to the maintenance of trust.

Where access to data has been approved, delays may still be experienced by data applicants in being able to access it from local organisations. Research may benefit therefore from measures to identify and address any unnecessary barriers or duplication that might arise from local consideration of issues already addressed by CAG.

Requests for access to health data held by NHS Digital is considered by the Independent Group Advising on the release of Data (IGARD). IGARD is an independent committee that reviews requests for NHS Digital data in line with its terms of reference. Similarly, to the aforementioned groups, IGARD should not be thought of as a barrier. These oversight groups are a crucial safeguard, as their role in the data access process provides vital reassurances to the public about the safety of their data in a research context. They give people peace of mind that there are stringent checks in place, and that data isn't a 'free for all'.

With the forthcoming merger of NHS England, NHSX, and NHS Digital, it will be important to ensure that good governance and independent oversight is maintained rather than diminished. It is vital that safeguards are respected, and standards are maintained, whether requests for data come from inside or outside of the new organisation.

Trusted Research Environments

A significant amount of work is currently underway within health and care to develop trusted research environments / secure research environments (referred to as TREs) as a safer mechanism for providing researchers with access to health and care data. The aim, which the NDG supports, is to evolve current data stewardship practice from data disseminations (sending extracts of the data off site to the user) to data access. This is a positive change that should reassure the public. Existing citizens' jury

research demonstrates that the public puts more trust in data access through software platforms such as OpenSAFELY, where those accessing data cannot make additional copies.

Alongside the technical development of these data access platforms, rigorous governance frameworks are also essential – so that sound decisions are made about who can access data within a TRE and on what basis. Given the potential scale of TREs, they offer a good opportunity to invite public/lay representatives onto the groups responsible for making those decisions.

National Data Guardian's response to recent government consultations

The call for evidence asked for respondents to state 'The extent to which data issues are appropriately addressed by the government's National Data Strategy, its draft strategy, data saves lives: reshaping health and social care with data, and its consultation Data: a new direction'.

Comprehensive feedback from the NDG to 'Data: a new direction' and 'Data Saves Lives' can be found in the links below:

National Data Guardian feedback on 'Data: a new direction': proposed government reforms to the UK data protection regime - GOV.UK (www.gov.uk)

National Data Guardian feedback on DHSC's draft data strategy: 'Data Saves Lives: Reshaping health and social care with data'. - GOV.UK (www.gov.uk)

Ethical data use

The ethics underpinning the use and sharing of individuals' data in health and care contexts

As previously mentioned in this response, health and social care data is often uniquely sensitive in nature, and patients want reassurance that it is being protected and used appropriately. When patients and service users provide their information to a care professional, they cannot be expected to know all the other ways in which it might be used. Legal frameworks

govern the use of data, but the need for strong ethical underpinnings that build on the protections provided by the law are also important.

Findings from public attitudes research projects demonstrate that many people feel strongly about their data being shared for secondary uses; they may have a variety of questions about it, such as:

- how is the data about me being used, where does it go and who sees it?
- do I have ownership over data about me, even if that data is in pseudonymised or anonymous form?
- what can I / should I be able to opt out of when it comes to the different uses to which data about me is being used?
- who is going to benefit when my information is used for reasons other than my own personal care?
- how does the NHS make decisions about who gets access to our data, and are private companies benefiting financially using data about me?
- does the NHS get a good deal when private companies access our data, or does it loose out longer-term?

Many of these are ethical questions, as ethics is concerned with what is good for individuals and society. And so to act ethically, we must engage with the public about such questions to understand what they believe is ethical, then create policies and make decisions that are influenced and guided by their views.

Transparency, reasonable expectations, and the importance of no surprises

Telling people how data about them is used is a cornerstone of public trust. This includes providing information that is accessible, concise and easy to comprehend, so that people from all walks of life, and of all ages and abilities, can easily understand how data about them might be used. Where this expectation has been set, organisations should then only use the data within the boundaries of that understanding. The requirement for transparency is also essential to any proposed use of data beyond health and care, including across government.

The National Data Guardian recently added an 8th Caldicott Principle to underline the importance of these considerations. The new principle's purpose is to make clear that patient and service users' reasonable expectations must be considered and informed when confidential

information is used, to ensure ‘no surprises’ about the handling or sharing of their data:

Principle 8: Inform patients and service users about how their confidential information is used

A range of steps should be taken to ensure no surprises for patients and service users, so they can have clear expectations about how and why their confidential information is used, and what choices they have about this. These steps will vary depending on the use: as a minimum, this should include providing accessible, relevant and appropriate information - in some cases, greater engagement will be required.

The importance of public benefit to data sharing

Research has demonstrated that for people to accept the use of their data for secondary purposes, they need reassurance that those purposes will deliver benefit back to individuals and society.

The NDG undertook a public dialogue project to understand how people perceive ‘public benefit’ when it comes to secondary uses of health and social care data: Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data - GOV.UK (www.gov.uk).

The NDG is currently using this public dialogue to develop content that will provide some clarity on the definition and evaluation of public benefit.

Artificial Intelligence and data driven innovation

Ethical considerations arise more acutely in new and emerging areas of innovation such as artificial intelligence, given the potential scale of both benefits and risks, including those arguably unique to the technology itself. In broad terms, the considerations remain the same, including a systematic evaluation of potential public benefit, ensuring access to confidential patient information is minimised, and that all use is underwritten by the principles of authentic public engagement and transparency.

There are, however, some distinct considerations. Firstly, and perhaps most fundamentally, there needs to be recognition of, and deep engagement with, the extent to which human decision-making power will determine AI impact.

In terms of privacy, developers need to proceed with extreme caution when handling any identifiable health and care data. Where possible, anonymous

data should be used and developments in synthetic data use are encouraging. Certain types of data may also require specific privacy considerations, for example with imaging there is no consensus yet on what can be considered a truly anonymous head scan. As with cybersecurity, privacy protections will need to continue to develop in tandem with emerging risks. Boundaries between different public sector organisations and government departments will also need to be consciously maintained; access for one purpose within health and care might lead to unanticipated findings or uses outside, which potentially could undermine trust in a confidential health and care system if data is used in ways the public does not expect or necessarily support.

As the horizons of what it's possible to do with health and care data expand, we need to ensure that our understanding of what people want, expect and feel comfortable with, evolves with it.

In summary

Improving the collection and use of health and care data will unlock further opportunities to benefit people's health and wellbeing, and to ensure the sustainability of the health and care system. Achieving this will require a deep understanding of the uniquely sensitive, confidential nature of health and care data, provided by patients and service users within the context of a reciprocal relationship with the health and care system that is based on trust. The public should not be asked to simply trust how their data is subsequently used. Rather the onus should be on the system to demonstrate its trustworthiness through a commitment to good governance, engagement and transparency, data security, the provision of authentic public choice, and ensuring that the public are represented through involvement in decision making.