



Volume 11 Issue 1



RESEARCH  
ARTICLE



OPEN  
ACCESS



PEER  
REVIEWED

## Towards responsible, lawful and ethical data processing: patient data in the UK

**Tess Johnson** *University of Oxford*

**Konrad Kollnig** *University of Oxford*

**Pierre Dewitte** *KU Leuven*

**DOI:** <https://doi.org/10.14763/2022.1.1638>

**Published:** 18 March 2022

**Received:** 14 September 2021 **Accepted:** 5 March 2022

**Competing Interests:** The author has declared that no competing interests exist that have influenced the text.

**Licence:** This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. <https://creativecommons.org/licenses/by/3.0/de/deed.en>  
Copyright remains with the author(s).

**Citation:** Johnson, T., Kollnig, K., & Dewitte, P. (2022). Towards responsible, lawful and ethical data processing: patient data in the UK. *Internet Policy Review*, 11(1). <https://doi.org/10.14763/2022.1.1638>

**Keywords:** Ethics, Data protection, Health, Big data, United Kingdom

**Abstract:** In May 2021, the UK National Health Service (NHS) proposed a scheme—called General Practice Data for Planning Research (GPDPR)—for sharing patients' data. Under that system, a patient who does not wish to participate must actively opt out of their data being shared with third parties for research and other purposes. In this paper, we analyse the lessons that can be learned for the responsible and ethical governance of health data from the NHS' new scheme. More specifically, we explore the extent to which the opt-out within the planned scheme complies with the requirements stemming from the General Data Protection Regulation (GDPR), particularly in relation to the principles of lawfulness and transparency. We then evaluate, from an ethical perspective, this opt-out 'nudge' and whether it is sufficiently resistible, reversible, and has appropriate goals. In light of the above, we then propose improvements for the scheme's legal and ethical acceptability.

## Introduction

The health care system is becoming increasingly digitised, resulting in a vast and growing amount of readily available patient data. This data can be extraordinarily valuable for research and other purposes in the public interest. For instance, the RECOVERY trial in the UK (Nuffield Department of Population Health, 2021) discovered new life-saving treatments for COVID-19 in a large-scale study with more than 40,000 participants, significantly reducing the lethality of the virus.

Yet, the use of health data raises many legal and ethical questions. If a healthcare system were to share patients' health data with commercial actors against their explicitly expressed wishes, this would be uncontroversially immoral and illegal. However, both the law and ethical guidance remain less certain on borderline cases—say, where data is shared for both commercial *and* research reasons, and where patients' consent is presumed when they fail to actively refuse a particular option ('opt-out systems'). In this paper, we consider a case study from this grey area. We investigate how legal and ethical analyses may, together, highlight problems with grey-area cases and provide guidance as to how changes to data sharing systems may make them legally and ethically acceptable.

Our case study is a recent opt-out data sharing policy proposed by the UK National Health Services (NHS)<sup>1</sup>, known as the General Practice Data for Planning Research system (NHS Digital, 2021a), but referred to as the NHS data scheme (NHS-DS) in this paper. The NHS-DS aims to share patients' health data for research purposes, and also with commercial actors. The planned system is similar to the one already in place for hospital data, and would extend this existing scheme to patient data from doctors' visits. Patients initially had six weeks (until 23 June 2021) to refuse access to their data, with the deadline later postponed (NHS Digital, 2021b) after public backlash. This backlash aligned with additional guidance on pre-conditions for deploying the NHS-DS, as follows (National Data Guardian, 2021):

1. The ability for patients to opt-out even after data collection
2. The creation of a 'Trusted Research Environment' for researchers to work with patient data without leaving NHS premises
3. A successful public campaign, 'explaining how data is used and patient choices'

1. The NHS is by no means as coherently structured as it may seem at first sight. Rather than being one central organisation, its functions are distributed across many organisations and across the UK. In the particular case of the NHS-DS, the organisation in charge is *NHS Digital*. Since all parts of the NHS coordinate with another and share the same public mandate to provide healthcare in the UK, we only refer to *the NHS* in our paper, for simplicity.

#### 4. Full clearance of the backlog of opt-outs received by GP practices

The NHS-DS provides a useful grey-area case, particularly regarding the legal and ethical design of public data institutions that enrol participants, and gather, guard and distribute their sensitive data. Both the new requirements put forward by the NHS already and further requirements may be needed to ensure ethically responsible and regulation-compliant data sharing.

There are strong arguments in favour of the use of health data for the public good, including the data held by the NHS. For instance, the UK RECOVERY trial cited above used samples of NHS patients' data (with their explicit consent) to identify a commonly used steroid, dexamethasone, as a viable treatment option for COVID-19 (Nuffield Department of Population Health, 2020). This new research knowledge and the resulting treatments provided significant public benefit. Aside from direct research benefits, funding is also a consideration in favour of data sharing. The consultancy EY estimates that the benefits derived from NHS health data could be up to 10bn GBP—per year (Ernst & Young, 2021). This amounts to 1/15 of the NHS' annual budget (Triggle, 2018), and could help increase the funding of the UK's main healthcare provider significantly. High numbers of sign-ups would be essential (Dash et al., 2019). If everyone refused to share their data as contributions to research knowledge and a sustainable public healthcare, no one would be helped. Thus, we have an initial argument for data-sharing systems. Indeed, it seems legitimate, too, to explore ways to increase the number of sign-ups, including opt-out strategies, whilst ensuring they are legal and ethical.

Unfortunately, those patients with extensive health records might both have the most useful health data for research purposes and also be the most fragile or ill individuals, already worse off than most. This raises particular concerns if sharing their health data risks making them worse off, still. This may be the case if release of their data leads to increased stigma or discrimination against them, either as individuals, or as members of a vulnerable population (Arias et al., 2015). As explored already in research on HIV/AIDS, sharing health data such as case reports concerning someone with a stigmatised disease can contribute to important research on new treatments and health surveillance to prevent further disease spread (Fairchild et al., 2007). However, given the stigma associated with the disease, the individuals in question may also have more reason than others to ensure their data remains shared only between them and their doctor. On a larger scale, harms to these disadvantaged groups that result from sharing their data present an important social justice concern in the age of big data and AI (boyd & Crawford, 2012; Binns et al., 2018), especially for those who believe it is wrong to impose

harms on those already worse off.

In the remainder of this paper, we first introduce the concept of data institutions (Section 1), and previous attempts in the UK to create data sharing agreements. We then assess the NHS-DS through the lenses of UK data protection law (Section 2) and nudge ethics (Section 3). Finally, finding that the NHS-DS as initially proposed requires changes to be acceptable, we suggest possible alterations that could render it an acceptable data sharing agreement (Section 4).

Our data collection mainly relied on the available information on the official website of the NHS-DS, especially the privacy policy and information on the opt-out process for patients. We additionally drew on relevant primary and secondary sources, including official statements by public and private bodies (including doctor and patient organisations), opinions of academics, as well as newspaper articles. We then used this data for further analysis, from both a legal and ethical perspective as explained in the corresponding sections.

## Background

The NHS-DS demonstrates one among many UK attempts to establish organisations for data stewardship in the health domain.

Already in 2006, the UK established the UK Biobank (2021), which has collected genetic information and biological samples from half a million UK citizens, and made this data available for research purposes under strict conditions. Interestingly, it has been possible for the UK Biobank to sign up a large number of participants *with their explicit consent* (UK Biobank, 2007). The Biobank can provide enormous benefits to research in the UK, since these participants provide active support to this project, such as participation in follow-up studies. The UK Biobank is monitored by the UK Biobank Board and the Ethics Advisory Council to ensure the ethical and legal stewardship of its data. Beyond health, other important data organisations based in the UK are the Secure Research Service (SRS) by the Office of National Statistics (2021) (making UK Census data available to accredited researchers), OpenCorporates (2021) (making company information available in the public domain), Open Banking Limited (2021) (establishing standards for inter-bank collaboration and data exchange), and HiLo (2021) (collating data for the maritime industry) (Open Data Institute, 2021). There is, then, evidence of many UK organisations that collect and use big data in a way that has not attracted public outcry, and is seen as conforming to norms of responsible stewardship.

Despite these successes, the UK has also run into significant challenges in establishing organisations for health data stewardship before. The most notable is the NHS' previous planned data-sharing system *care.data*, which aimed to make GP data available more widely (McKee, 2014). After more than a million people opted out, the policy was scrapped in 2016 (Temperton, 2016). Sterckx et al. (2015) argued that *care.data* failed to provide patients with real choice and enough transparency leading to a lack of trust in the project and its ultimate failure. In a separate case from 2015, the Royal Free London NHS Foundation Trust agreed to share the data of 1.5 million patients with DeepMind. This agreement was later found to violate UK data protection law (Denham, 2017; Basu and Guinchard, 2020), but only after data had already been shared and DeepMind had been acquired by Google. These are just two examples of a series of data sharing agreements that have been criticised for their lack of transparency, patient choice, and public consultation. They highlight the challenges in establishing agreements for responsible stewardship of data.

Motivated by the potential benefits for patients, the stewardship of health data is increasingly becoming an important function of the NHS itself, aiming to govern data relating to its patients and to make this data available to third parties. A key necessity for such organisations is ensuring participants' continued support in data-driven health research. One common way to ensure support is gaining participants' *meaningful* consent. However, gaining and maintaining consent, particularly via opt-out policies, can be challenging, both legally and ethically. The analysis we perform here, along with our suggestions for improvements, are broadly useful for guiding improvements in ethical and legal terms to data sharing agreements and the responsible implementation of data institutions.

## Implications under UK and EU data protection law

The GDPR, as implemented in the UK<sup>2</sup>, provides special protections for data that relates to individuals<sup>3</sup> ('personal data'). All UK organisations that determine the 'purposes' and the 'means' of such processing ('controllers'), including the NHS<sup>4</sup>,

2. While the UK has officially withdrawn from the European Union on 31 January 2020, meaning that the EU General Data Protection Regulation ('EU GDPR') does not apply domestically anymore, the text has nonetheless been brought into UK law as the UK General Data Protection Regulation ('UK GDPR') that coexists with the amended version of the Data Protection Act 2018 ('DPA'). This means that the core principles of the EU GDPR have outlived UK's adherence to the EU and are still relevant post-Brexit, as illustrated by the recent adequacy decision issued by the European Commission (2021).
3. More accurately, personal data is defined as 'any information relating to an identified or identifiable natural person' (Article 4(1) GDPR).

must comply with the GDPR, or face potentially high fines for non-compliance.

## The applicability of the GDPR

The data processed in the context of the NHS-DS is *not anonymised* since it can be linked to the patients by converting 'unique codes back to data that could directly identify patients in certain circumstances' (General Practice Data for Planning and Research, 2021). Such data is known as *pseudonymised data* (Article 4(5) GDPR) and means the NHS-DS still falls under the GDPR (Recital 26 GDPR; Article 29 Working Party, 2007, p. 18). Proper anonymisation would have allowed the NHS-DS to side-step the GDPR, but it is difficult to achieve and may impede the usefulness of the data for research (Article 29 Working Party, 2014a, p. 3). The NHS does, however, undertake some efforts to make the data less sensitive by, for instance, removing some personal attributes, including names and addresses from the records.

## The general principles governing the processing of personal data

Controllers, must comply with various rules and principles, including the obligation to ensure adequate transparency, to rely on one of the six lawful grounds listed in Article 6 GDPR, and to abide by all the other general principles listed in Article 5 GDPR (i.e. fairness, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality, and accountability). The GDPR also grants individuals whose personal data are being processed ('data subjects') a series of rights listed in Articles 15–23, such as the right of access or erasure. Of particular relevance in the case of the NHS-DS are the (i) general prohibition of the processing of health data, as well as (ii) the lighter regime for the processing of personal data 'for scientific research purposes', both of which are discussed below.

### A prohibition on the processing of health data

As a general principle, the GDPR prohibits the processing of 'special categories of personal data'. This includes personal 'data concerning health' (Article 9(1) GDPR). The datasets processed and shared in the context of the NHS-DS qualify as such. This prohibition can, however, be lifted by relying on one of the ten exemptions listed in Article 9(2) GDPR. In the case of the NHS-DS, three exemptions appear relevant.

4. In complex data processing situations, controllership is often shared with other entities ('joint controllership'). For simplicity, we assume that the NHS is the only data controller in this work. In particular, we do not consider the obligations of third parties that may get access to NHS data, and thereby might enter joint controllership with the NHS over patient data, but solely focus on the obligations of the NHS as the organisation currently being in control of this data.

The NHS could have gathered the data subject's explicit consent (Article 9(2)a GDPR). This would have proven complex, however. While the GDPR sets a high bar for what constitutes valid consent (Article 4(11) GDPR), an even higher bar of 'explicit' consent would be required to lift the prohibition around the use of personal health data (e.g. involving a written declaration from the data subject) (European Data Protection Board, 2020, para 93).

Another option would have been to justify that the processing is 'necessary for [...] scientific [...] research purposes' (Article 9(2)j GDPR). The use of that exemption also raises multiple difficulties. First, it would require the NHS to demonstrate that each processing activity is objectively 'necessary' for research purposes (i.e. that there is no less privacy-invasive way to do so<sup>5</sup>). Second, that exemption only applies to processing activities that actually pursue 'research' purposes, a notion the definition of which is nowhere to be found in the text of the GDPR (European Data Protection Supervisor, 2020, p. 9; Ducato, 2020, p. 3). Finally, the NHS would need to rely on a domestic law that specifies the details of the data processing and ensures adequate safeguards.

Instead, the NHS chose to rely on Article 9(2)g GDPR and argued that the sharing of patient data was 'necessary for reasons of substantial public interest', as explained in its transparency notice. Subject to the same legality requirement applicable to the research exemption detailed above, the NHS identified the General Practice Data for Planning and Research Directions 2021 as the appropriate piece of domestic legislation (itself implementing the Health and Social Care Act 2012).

It is worth noting that, beyond the requirement to identify an exemption from the prohibition of processing special categories of personal data (Article 9(1) GDPR), the controller must *also* rely on one of the six lawful grounds listed in Article 6(1) GDPR (Information Commissioner's Office, 2021, p. 87; Georgieva and Kuner, 2020, p. 376; Article 29 Working Party 2014b, point III.1.2). In this case, and since the choice of Article 9(2)g already required the NHS to identify a relevant piece of domestic legislation, it opted for 'legal obligation' (Article 6(2)c GDPR).

5. In that sense, a parallel can be made with the necessity test detailed in the Article 29 Working Party Opinion 06/2014 on the notion of legitimate interests (Article 29 Working Party 2014b, point III.1.1.) and the European Data Protection Board Guidelines 2/2019 on the processing of personal data under Article 6(1)b GDPR (EDPB, 2019, point 2.4). It also stems from the first element of the checklist made available by the ICO under the section 'Special category data' (i.e. 'We have checked the processing of the special category data is necessary for the purpose we have identified and are satisfied there is no other reasonable and less intrusive way to achieve that purpose'). (Information Commissioner's Office, 2021, p. 87).

## A lighter data protection regime for scientific research purposes

Given the potential benefits of the processing of personal data for scientific research purposes, the GDPR has laid down a *lighter data protection regime*. It allows some rules of the GDPR to be relaxed, provided that the research activities meet certain criteria. First, as hinted above, these activities need to pursue actual 'scientific research' purposes. While there is no strict definition of that notion, Recital 159 GDPR<sup>6</sup> specifies that it 'should be interpreted in a broad manner including, for example, technological development and demonstration, fundamental research, applied research and privately funded research'. Thus, research does not have to be non-commercial to qualify for that lighter regime, even though individuals' concerns might substantially vary depending on its public or private nature (Verheenen et al., 2020, p. 38). As a result, the fact that the NHS-DS envisions health data transfer for commercial purposes does not, *per se*, disqualify it from the lighter regime.

Second, and in order to benefit from that lighter regime, controllers must implement appropriate safeguards to protect data subjects' rights and freedoms. This includes an obligation to use pseudonymisation, or even anonymisation, if possible (Article 89(1) GDPR). The NHS and any third parties must therefore assess whether the research can yield conclusive results with anonymised or pseudonymised datasets. According to its GP Practice Privacy Notice (2021), the NHS-DS does indeed make use of this lighter regime, and thus needs to implement the above safeguards.

Research that meets the above criteria benefits from exemptions from some GDPR rules. First, the principle of purpose limitation is relaxed (Article 5(1)b and Recital 50 GDPR; European Data Protection Supervisor, 2020, p. 23). As a result, the NHS-DS does not have to rely on a different lawful ground than the one specified for the collection by the GPs. So is the principle of storage minimisation. Personal data can therefore be stored in a form which permits the identification of the data subjects for longer than necessary for the purposes for which they were initially collected (Article 5(1)e GDPR)—though not indefinitely. Third, data subjects are deprived from the possibility to exert their right to erasure once the research has already been concluded and published (Article 17(3)d GDPR).

6. The text of the GDPR comes with as many as 173 *Recitals* that explain in more detail what is meant by the Articles of the law. While only the Articles are legally binding, the *Recitals* are important to resolve any ambiguities in the application of the law in practice and in court.

## Further data protection obligations

Regardless of the lighter regime outlined above, the NHS-DS still involves the processing of personal data and, as a result, the NHS must comply with all the principles and rules outlined in the GDPR. More specifically, it will have to ensure adequate transparency by providing data subjects with the information listed in Article 14 GDPR ‘in a concise, transparent, intelligible and easily accessible form, using clear and plain language’ (Article 12(1) GDPR).

Despite these obligations, the planned data sharing system has already been criticised for not adequately informing the public of plans for patient data and how to refuse to participate (Boiten, 2021). Public debate has been limited, and information was not initially sent to households. This contrasts with the previous care.data system, wherein individual households received information on the benefits of this system (although, even then, not details on how to opt out) (Solon, 2014).

While transparency and consent are often in the spotlight of the public debate around data sharing, there are other concerns. According to the *data minimisation* principle in the GDPR, the NHS-DS must limit its processing to the personal data that are objectively necessary to achieve the purposes it strives to achieve (Article 5(1)c GDPR). Data might also be stored centrally for an unspecified time for the NHS-DS and third parties, which is in conflict with the principle of *storage limitation*. It is also not clear how the NHS-DS implements the principle of *purpose limitation*, which essentially requires the definition of *specific* objectives for data use from the onset. Beyond these measures, trust is required in the third parties, that they will not to use data for purposes other than those initially agreed upon. *Integrity and confidentiality* (Article 5(1)f GDPR) and *security* (Article 32 GDPR) also play a crucial role when it comes to the processing and storing of large amounts of personal data at the national level. Protecting centrally stored data against cyber-attacks and accidental breaches is difficult. This is even more complex for data stored with third parties.

Our legal analysis underlines that robust measures must be put in place to address the risks arising from large-scale data processing of sensitive personal data, and the sharing of such data with third-party organisations. There are also some ethical issues that roughly align with the legal problems discussed above. In case both perspectives show the scheme to be unacceptable, there is increased justification to alter it.

## Ethical aspects

The NHS-DS uses an opt-out feature to sign up participants, sometimes called 'presumed consent' (Prabhu, 2019; Sterckx et. al, 2015) or a 'default nudge' (Schmidt & Engelen, 2020) in the ethics literature. In order to not perform the desired behaviour of sharing their data, participants must actively refuse the NHS-DS. The aim is to gain more individuals' agreement to share their health data for certain purposes.

According to models of consent, individuals have agreed to an action via presumed consent if they have been adequately informed of a proposed action, and fail to refuse it. The question of whether opt-out systems in fact adequately ensure participants' consent has been explored effectively in the context of organ donation after death (Miller, Currie & O'Carroll, 2019; Prabhu, 2019).

More recently, opt-out systems have also been discussed in the behavioural economics and ethics literature on 'nudges'. Nudges are 'any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives' (Thaler & Sunstein, 2009, p. 6). Default nudges such as opt-out systems for data sharing encourage the individual to share their data by making it automatic that they do so, rather than appealing to them via their rational faculties to change the behaviour. This generally results in more of the desired behaviour in a population, compared to under a no-default/forced choice or opt-in system of choice (Starr, 2000). Often, nudges are 'for good', aiming to benefit the individual or a community, and are implemented in ways that are considered ethically acceptable in terms of the range of choice they offer, their desired aims, and how obvious they are to an individual (Lades & Delaney, 2020). Indeed, nudges for good have been used in many public health programmes. For example, some anti-smoking campaigns mandate shocking images on cigarette packets which may nudge people against smoking, causing a so-called 'affective primer effect'. This works by altering an individual's emotional state before or as they make a behavioural decision (such as whether to smoke a cigarette), and can influence their decision (Friis et al., 2017). In the smoking case, the negative emotional effect of the shocking image on the packet may influence the individual not to have that cigarette

Opt-out systems in public health can also be an important nudge for good. Indeed, it may be important for the health of UK residents that the NHS shares data with research institutions, for the development of treatments and preventive measures (Sterckx et. al, 2015). However, opt-out data sharing may also be ethically prob-

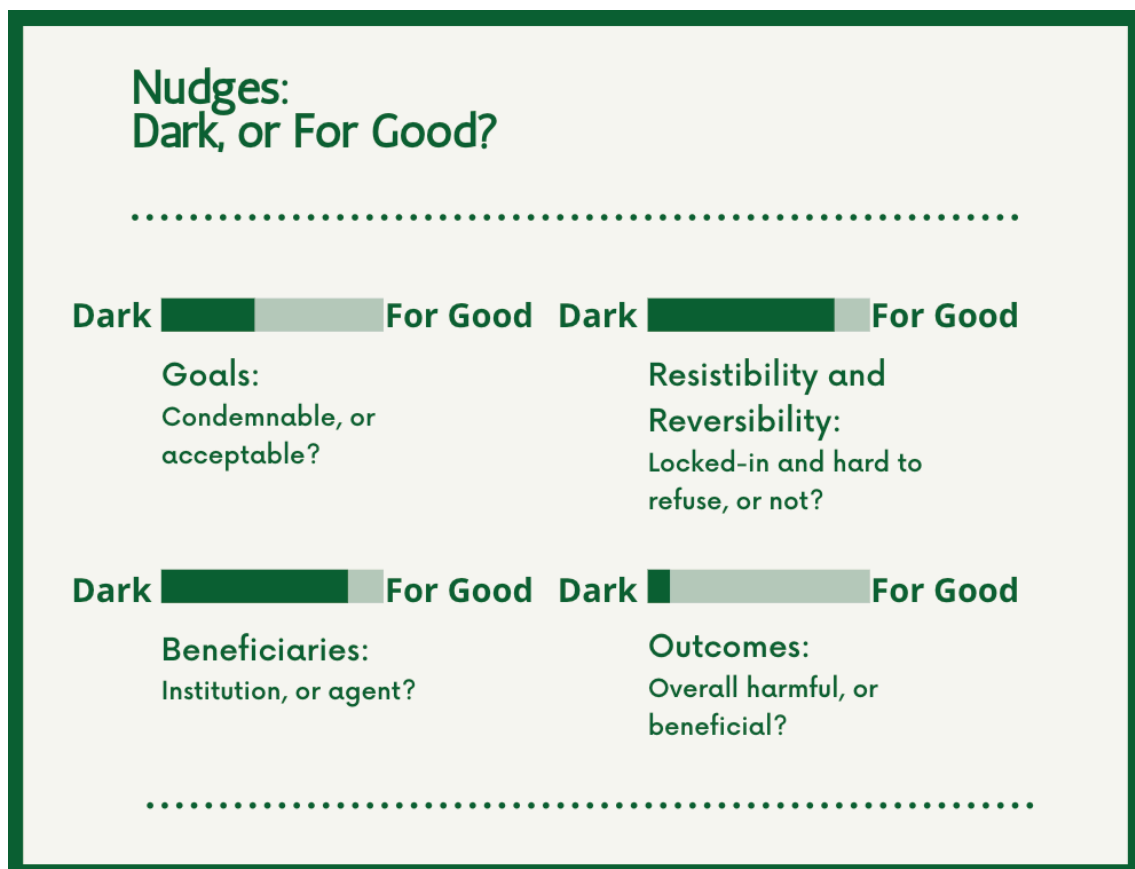
lematic, if this data is used commercially in ways that may undermine patients' interests. Furthermore, it may be ethically problematic if patients' choices whether to share their data are manipulated. Nudges that have these (and other) characteristics are sometimes termed 'dark' nudges or 'sludges' (Thaler, 2018). When it comes to sharing health data, we want to avoid exposing medically vulnerable people to uses of their data that they may not want, or that may violate their privacy, without giving them adequate opportunity to refuse (Arias et al., 2015; boyd & Crawford, 2012). For instance, commercial actors who have access to patients' data may at some point include health insurance companies (although this isn't currently the plan for data from the NHS-DS (NHS Digital, 2021c)). When insurers have information on patients' health status, less insurable or more costly prospective patients may find it difficult or expensive to buy additional health insurance, because they have shared their health data. Another concern is reciprocity. If commercial actors may benefit from access to health data, there ought to be similar benefits for patients, research participants and their communities, in exchange—as is, increasingly, demanded by the populations themselves (Merson et al., 2015). While using health data for research pays back patients for their data contributions when new treatments are developed and implemented in that community, this is not as clearly the case for data shared with commercial third parties.

## **Dark nudges and irresistibility**

Do opt-out systems for data sharing like the NHS-DS constitute dark nudges?

The literature on dark nudges has been particularly well-developed in business ethics. A new scalar model of nudges has been proposed, to distinguish between dark nudges and nudges for good, according to where they fall on these scales (BVA Nudge Unit, 2018). While these have not been well-explored in the academic literature, they are designed by the BVA Nudge Unit, a respected behavioural research institution, to provide a means for businesses to evaluate proposed nudges. The scales assess the goals, beneficiaries, outcomes, and resistibility and reversibility of a nudge, among others.

Dark nudges are on the left of each of these scales, adapted from the BVA Nudge Unit's original in Figure 1:



**FIGURE 1:**The scalar model of nudges, ranging from dark nudges to nudges for good on several separate scales, with no particular intervention represented.

Being *insufficiently resistible* is one feature of a dark nudge, according to the scalar model above. Default nudges harness people's tendency to stick to the status quo, their inertia. They should do so in a way that still allows individuals to change their minds and behaviour easily, if a nudge is to be counted as 'for good'. More formally, this has been stated as the requirement that nudges be *substantially non-controlling*: 'A's influence to get B to  $\phi$  is substantially noncontrolling when B could easily not  $\phi$  if she did not want to  $\phi$ ' (Saghai, 2013, p. 488).

Another way that dark nudges may be ethically problematic is by being non-transparent, and locking people into their choices. If people are not informed 1) that they're making a choice; 2) of how they can refuse the nudge, both now and at a later point, then resisting and reversing the nudge becomes more difficult, and the risk that the nudge wrongfully interferes with the individual's choice in some way may become greater, as argued in the context of non-consensual neurointerventions (Douglas, 2022).

Let us consider the NHS-DS using the scalar model, then. The system has been

poorly publicised, and the ways to opt out require multiple steps on the online website to determine where and how to send the opt-out form. In fact, the authors' process of trying to determine how many clicks it takes to get from the NHS homepage to the opt-out form for the NHS-DS, we found ourselves being redirected to the 'National Data Opt-out', instead of the necessary 'Type 1 Opt-out' to stop data collection completely (see Figure 2). Further, the form is hardcopy, and no online opt-out exists at the time of writing. This lack of transparency and easy refusal pushes the NHS-DS as it stands closer to a 'dark nudge', at least on the scale of resistibility and reversibility. While the outcomes scale is difficult to assess in advance, the next subsection considers where the NHS-DS' beneficiaries and goals place it on the spectrum from dark nudge to nudge for good.

---

## Opting out

If you don't want your identifiable patient data to be shared for purposes except for your own care, you can opt-out by registering a [Type 1 Opt-out](#) or a [National Data Opt-out](#), or both. These opt-outs are different and they are explained in more detail below. Your individual care will not be affected if you opt-out using either option.

---

**FIGURE 2:** Finding the necessary form to opt-out from collecting GP patient data on the NHS-DS website is hard for individuals, and they must submit the form to their GP practice in print. In the above picture, both links lead to the National Data Opt-out, and not to the necessary Type 1 Opt-out form—even ignoring the ambiguity of both terms (NHS Digital, 2021a).

---

## Dark nudges and condemnable goals or inappropriate beneficiaries

Dark nudges are often characterised as having harmful overall goals—either for the individual, or their community. Consider these as the antitheses to nudges that benefit either the individual—'pro-self' nudges (Hagman et al. 2015)—or a community—'pro-social' nudges. Even then, pro-social nudges, in influencing an individual to benefit others, may be at greater risk of becoming dark nudges, compared to pro-self nudges, if they treat the individual as an instrument to benefit others. Indeed, research by William Hagman et al. (2015) found that pro-social nudges are significantly less acceptable to members of the US and Swedish publics, than pro-self nudges.

Does the NHS-DS fit the bill for a pro-social nudge for good, sharing health data to contribute to health research? Or does it constitute a dark nudge on this scale? Theoretically, it is simply offering people the choice to benefit others (and, in some cases, themselves) by sharing their data for research purposes (although, also commercial ones). We have seen the benefits of data sharing in the UK RECOVERY tri-

als (Nuffield Department of Population Health, 2021). However, research is not the only intention behind the NHS-DS. It also aims to share the data with third parties, including commercial actors. Commercial interests in patient data may seem to constitute ‘illicit ends’ (Schmidt & Engelen, 2020) or harmful overall goals of a nudge, pushing it toward the dark nudge end of the scale. Indeed, while we lack the scope to explore this issue fully here, the NHS itself has a financial interest in patients’ data, which further complicates the issue of its being the agency to nudge patients in the first place.

Then again, there may be good reasons even for this additional, commercial sharing of data. As patient needs increase and funding fails to fill the breach, the NHS has turned increasingly to private ventures to support the system. One way that the NHS can continue its service is by selling some of its invaluable patient data (Horn & Kerasidou, 2020; Carter et al., 2015). Furthermore, data sharing may actually express solidarity within the healthcare system, and promote trust in the system, constituting a benefit of patients’ decisions to share their data (Horn & Kerasidou, 2020). If one of the goals of commercial data sharing is to promote the continuity of, trust in, and solidarity with the NHS, this may be a laudable goal.

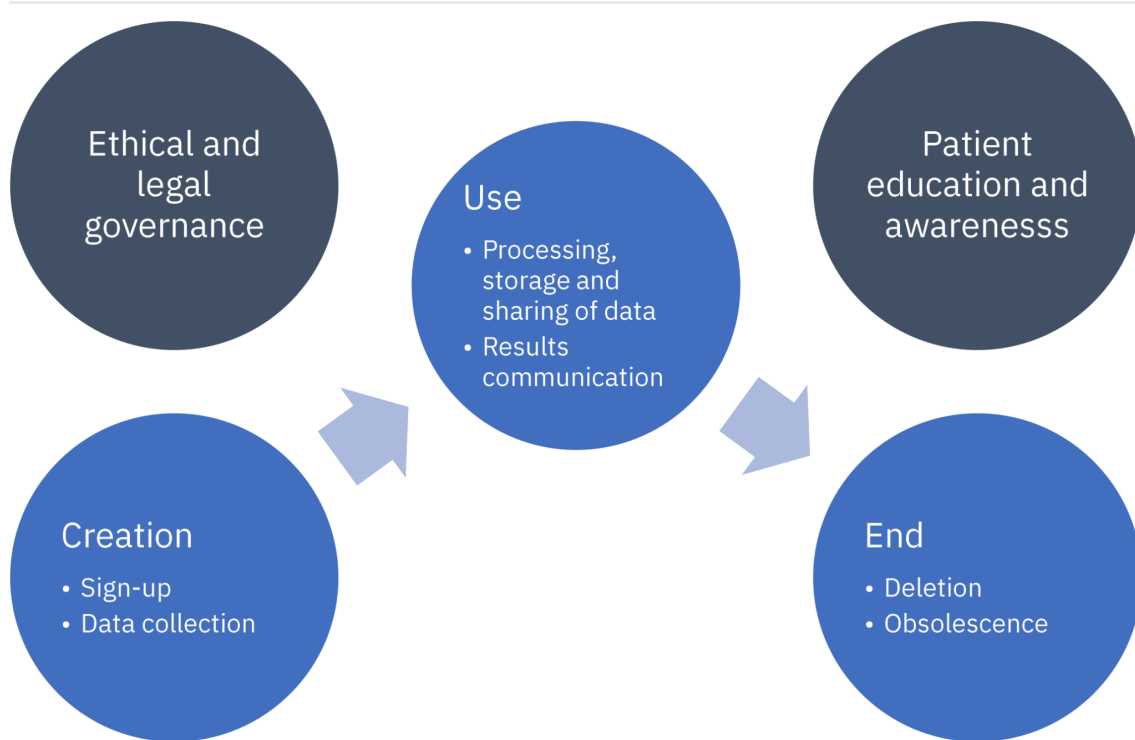
The NHS, then, may also be an appropriate beneficiary of data sharing, indirectly benefiting as a public institution by compensation for third party access to data. Yet, the point remains that it is not the only beneficiary. Commercial actors are also beneficiaries. Whether they are appropriate ones may depend on characteristics of the companies themselves, and on their prospective uses of the data. These uses remain unclear, as does the access these companies may have in the future to patient data. Will patients be adequately compensated, fulfilling ethical requirements for reciprocity, if their data is very helpful for, say, targeted advertising? Might the data be put to use in a way that undermines patients’ interests, say by health insurers, to increase premiums? Nudges that have commercial enterprises as beneficiaries instead of patient communities risk shifting toward dark nudges. While the division of community vs company as beneficiaries is unclear, the NHS-DS certainly has both.

Rather than a nudge for good, the NHS-DS may constitute a dark nudge, given its low resistibility and transparency, and its mixed goals and beneficiaries. In that case, it seems ethically unacceptable, such that both legal and ethical judgements of the proposed system align. However, recent updates to the plan for implementing the NHS-DS may help push the scheme closer to a nudge for good, as may further changes. In that case, even if the opt-out system is maintained, the NHS-DS may be ethically acceptable. This judgement may, to an extent, align with the lim-

ited acceptability of opt-out systems for data collection according to GDPR law, even without explicit consent, according to the exemption discussed above. However, the NHS-DS as it stands is not considered acceptable either from a legal or ethical perspective. This may, first, expose ways in which the NHS-DS should be changed to make it acceptable. Furthermore, if there are any remaining mis-alignments between legal and ethical judgements of a new and improved NHS-DS, this may indicate ways in which GDPR law itself could be improved, as is currently attempted by the UK government (UK Department for Digital, Culture, Media & Sport, 2021). In the next section, we explore how some legal and ethical issues we have identified can be avoided, while maintaining the data sharing system in a similar form.

## **A more responsible data lifecycle**

We now turn to the implications of our above analysis, by making a range of suggestions on how to make data sharing schemes more ethical and legal, and ultimately more accepted by the general public. The suggestions we make here occur at several stages of the *data lifecycle*, and intersect with the key problems we have identified with the NHS-DS. These are 1) creation, consisting of the initial collection of data (particularly consent to data sharing) and putting safeguards in place; 2) use, consisting of the processing, storage and sharing of data and communication of results; and 3) end, data deletion or obsolescence. Our suggestions cover each of these areas, contributing to the overall governance of the data use (Figure 3). An overarching challenge relating to these three stages in the data lifecycle is patient education and awareness.



**FIGURE 3:** Summary of the main lessons learned from the NHS-DS case study, particularly regarding the data lifecycle and the governance of such a system.

### Creation: Sign-up and data collection

If an opt-out system is to pass legal and ethical muster, the process begins with ensuring that patients have agreed to—or at least could have disagreed but have not—participating in a data sharing agreement. Even reasonable data uses must be able to be refused by patients, even if anonymised data is used (Sterckx et. al, 2015). Second, wherever commercial actors may use patient data, we cannot assume that data use will be in patients’ best interests, and that they will agree to it.

To mitigate these issues, one strategy is to publish a full data protection impact assessment (DPIA) (not only a baseline version, as suggested by the NHS’ proposed system (NHS Digital, 2021d)). While currently not a legal obligation, it is nonetheless good practice to do so, especially when it comes to high-risk processing of special category data, and to personal data held by large and influential, publicly-funded organisations. This must be published *before* data processing takes place, in order to be of use to potentially non-consenting patients, and in alignment with existing requirements (GDPR, 2016, Recital 61). The DPIA must not only be published, but must have at least a version that is easily accessible and readable for a public audience (whilst the main document may be aimed at specialists).

Once patients have (if they choose) learned the details of the NHS-DS, it must be

easy for them to refuse to share their data, either, say, by (e)mailing a refusal, or clicking through a webform. That is, the nudge toward data sharing that patients experience must be sufficiently resistible. Such webforms are not mechanisms for adequate resistibility where they are hidden multiple clicks away or use non-intuitive wording. Finally, patients must be given adequate time to opt out. The NHS has already acknowledged this, by extending the time frame until the introduction of the NHS-DS. However, this opt-out still relies on paper forms, creating an unnecessary burden on patients wishing to refuse.

### **Use: Processing, storage and sharing of data, and results communication**

There are many techniques that can be used for better processing, storage and sharing of data beyond the scope of this article. As a general principle, data stored and storage time should be minimised, a central point of failure should be avoided (e.g. through decentralisation). This is particularly relevant in the context of cyber-attacks, which put any centralised system (and even those systems that enable the quick and easy access to decentralised information) at great risk.

When making data available to third-party organisations, *patient data should ideally never leave NHS premises*, and instead be kept in dedicated and secure computing environments. The NHS has already promised that, for now, data shall not leave NHS premises unless patients explicitly consent and instead be handled in a 'Trusted Research Environment' (NHS Digital, 2021b), though it remains unclear whether this might be changed in the future. This promise might be motivated in part by the negative experiences from the unlawful sharing of patients' health data with Google DeepMind.

If personal data must leave the NHS (and it may never need to), *effective privacy-preserving methods* (such as homomorphic encryption and secure multiparty computation (Scheibner et al., 2021)) must be applied. If possible, synthetic data (i.e. the generation of data that looks like real data, but is in fact generated through a statistical process, e.g. a computer simulation) should be used instead of real data (though some privacy issues remain with synthetic data, as well). As highlighted by our legal analysis, the NHS and any third parties must assess whether any research can yield conclusive results with anonymised or pseudonymised datasets, and must do so if possible.

While heightened technical protection and state-of-the-art measures come with a certain learning curve (Agrawal et al., 2021), education provided by the NHS to

practitioners could serve as *important leverage to propagate best practices* around ethical data use across the industry. Importantly, the use of state-of-the-art technologies for data protection can offer patients more control over their data if they agree to share it, including flexibility to opt out of various particular uses in the future.

## **End: Deletion and obsolescence of data**

As highlighted by our legal analysis, *patients should be able to opt out, at any time, even after the initial data collection*. If they choose to, their data should be excluded from the NHS databases as well as research in progress, but not from already concluded and published research. Furthermore, the principle of storage limitation under the GDPR limits the duration of data storage to what is objectively necessary.

## **Independent ethical and legal data governance**

It is important to have the right governance mechanisms in place to ensure ethical and legal processing of patient data, by both the NHS and third-party organisations. Especially in systems using nudges to recruit participants, data sharing must have appropriate goals and beneficiaries, to constitute a nudge for good. In health-care settings, for example, a primary goal should be to contribute to useful health research. Even if data is used to improve the funding situation of the NHS, the primary beneficiaries ought to be the patient communities. One way to protect these goals would be to have an independent organisation approve uses of the data. In particular, such an organisation should be independent from the NHS, with no interests relating to the use of patients' data (and especially not of financial nature). For the NHS-DS, this is proposed through the Independent Group Advising on the Release of Data (IGARD, formerly 'Data Access Advisory Group'), but also through engagement with the British Medical Association, Royal College of GPs and the National Data Guardian. It is important to ensure that the members of the independent organisation have sufficient expertise in all relevant disciplines (including medicine, law, ethics, technology). Similar principles apply to other data sharing institutions and the use of independent organisations (O'Hara, 2019). This will help to prevent mission creep, and ensure adequate oversight.

## **Ensuring patient education and awareness**

Beyond the three aspects that directly relate to the NHS-DS, it is important to ensure patients' adequate education on particular data sharing schemes, as both the donors of data, and the ultimate beneficiaries of data sharing. Participation in the

process and increased education builds trust and confidence in data sharing schemes, ensuring their sustainability, and that they continue to be informed by patient-oriented goals (National Academies of Sciences, Engineering, and Medicine, 2015; Sterckx et al., 2015). The data held by the NHS bears enormous potential, but the use of this data will not come without risks. While these risks should be listed in the DPIA, not every patient will be able to study this document in depth. It is therefore important to inform the public of the risks, benefits and choices attached to processing their health data, widely and in an intelligible manner—for instance, via public campaigns and schools. Educating patients about the benefits of data sharing (especially people with critical diseases) and addressing their privacy concerns could be important for achieving the aims behind the NHS-DS.

## Conclusions

Health data sharing schemes, such as the NHS-DS, bear great potential. There is, however, the need for balance between the rights of individuals and the interests of public institutions, both in legal and ethical terms. In particular, the use of an opt-out scheme and the sharing of patient data with commercial organisations can be acceptable, but strong safeguards must be in place. These are designed to protect medically vulnerable people from invasive data practices, and to avoid limiting the beneficial outcomes of data sharing agreements to only a small group of individuals (e.g. the shareholders of US tech companies). The necessary safeguards include the adequate informing of citizens, and easy resistibility and reversibility of the decision even after data has been collected. As it stands, the NHS-DS seems closer to a dark nudge than a nudge for good, although the recently proposed changes by the UK government might help improve the system.

In order to make the decision to opt out accessible, the system should be available both online and in physical form. When handling health data, it is also essential to use existing, state-of-the-art privacy protections, including the use of truly anonymised or synthetic data and dedicated computing environments where possible, rather than ever making real patient data available to third parties—at least without the necessary protections to prevent data misuse now and in the future. More generally, given the conflicts of interest of both third-party organisations as well as the NHS, a capable, independent oversight board is an absolute necessity, in light of swiftly changing legal and ethical expectations, and to avoid mission creep beyond current plans and promises.

Another major risk arises from the planned and necessary centralisation of sensi-

tive patient data in the NHS-DS. Over the past years, data breaches have been accelerating, and included numerous high-profile organisations such as the UK Parliament in 2017, British Airways in 2018 and Facebook in 2019. Large-scale data breaches are almost inevitable over the long lifespan of IT systems in the health sector. If a large data breach of NHS patient data were to occur and such data could be traced back to individuals, this could not only significantly harm the individuals affected, but also public trust in the NHS and other public organisations.

Instead of relying on legal mechanisms to avoid having to gain consent (such as public interest in data sharing), part of establishing a trust and solidarity in data sharing systems may be the publication of plans to institute an opt-out policy—way ahead of time. This is especially the case where trust is lacking, as the NHS' care.data scandal arguably demonstrated (Solon, 2014). If people are given information and choice, they may be more likely to support the system. However, if consent is wrongly presumed, inadequately informed, given without capacity, or inadequately voluntary, then the choice can make patients feel betrayed. That is not to say that opt-out data sharing agreements are always wrong. For many data stewards like the NHS, opt-in consent is neither feasible, nor necessary, because the system must deal with millions of people. However, a better opt-out system might improve trust and better constitute a nudge for good.

The lessons learned from our analysis are not just relevant for a UK context. Although our ethical and legal discussion employs a UK perspective, it is relevant more broadly. The legal framework of our analysis, the GDPR, is currently the same within the European Union. Beyond Europe, many countries have been introducing laws that are structurally and methodologically similar to the GDPR—most recently, China, with its Personal Information Protection Law (Kollnig et al., 2021). Our analysis might be especially relevant for countries with populations that are sceptical about public institutions, or that tend to be concerned about issues relating to data protection.

As our methods for large-scale data analysis will continue to evolve, we will, too, need to continue the debate around what data practices are permissible and how to implement them in practice while ensuring the trust and safety of those affected. Our paper hopes to make a contribution to this ongoing debate, but much more future research will be needed to tackle the current and emerging challenges in the use of health data.

---

## References

Agrawal, N., Binns, R., Van Kleek, M., Laine, K., & Shadbolt, N. (2021). Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3411764.3445677>

Arias, J. J., Pham-Kanter, G., Gonzalez, R., & Campbell, E. G. (2015). Trust, vulnerable populations, and genetic data sharing. *Journal of Law and the Biosciences*, lsv044. <https://doi.org/10.1093/jlb/lsv044>

Article 29 Working Party. (2007a). *Guidelines on Personal data breach notification under Regulation 2016/679*. European Commission. <https://ec.europa.eu/newsroom/article29/redirection/document/49827>

Article 29 Working Party. (2007b). *Opinion 4/2007 on the concept of personal data*. European Commission. [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf)

Article 29 Working Party. (2014a). *Opinion 05/2014 on Anonymisation Techniques*. European Commission. [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf)

Article 29 Working Party. (2014b). *Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC*. European Commission. [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf)

Article 29 Working Party. (2017). *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is “likely to result in a high risk*. European Commission. [https://ec.europa.eu/newsroom/article29/item-detail.cfm?item\\_id=611236](https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236)

Basu, S., & Guinchard, A. (2020). Restoring trust into the NHS: Promoting data protection as an ‘architecture of custody’ for the sharing of data in direct care. *International Journal of Law and Information Technology*, 28(3), 243–272. <https://doi.org/10.1093/ijlit/ea014>

Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). “It’s Reducing a Human Being to a Percentage”: Perceptions of Justice in Algorithmic Decisions. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3173574.3173951>

Boiten, E. (2021, July 27). NHS plan to share GP patient data postponed – but will new measures address concerns? *The Conversation*. <https://theconversation.com/nhs-plan-to-share-gp-patient-data-postponed-but-will-new-measures-address-concerns-165103>

boyd, danah, & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>

B.V.A. Nudge Unit. (2018). *Nudge, Dark Nudges, Sludge and Dealing with the Ethics of BE*. <https://bvannudgeunit.com/nudge-dark-nudge-sludge-dealing-with-ethics-be/>

Carter, P., Laurie, G. T., & Dixon-Woods, M. (2015). The social licence for research: Why *care.data* ran into trouble. *Journal of Medical Ethics*, 41(5), 404–409. <https://doi.org/10.1136/medethics-2014-102>

374

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: Management, analysis and future prospects. *Journal of Big Data*, 6(1), 54. <https://doi.org/10.1186/s40537-019-0217-0>

Denham, E. (2017). *RFA0627721 – provision of patient data to DeepMind*. <https://ico.org.uk/media/action-weve-taken/undertakings/2014353/undertaking-cover-letter-revised-04072017-to-first-person.pdf>

Department for Digital, Culture, Media, & Sport. (2021). *Data: A new direction* [Public consultation]. <https://www.gov.uk/government/consultations/data-a-new-direction>

Douglas, T. (2022). The Mere Substitution Defence of Nudging Works for Neurointerventions Too. *Journal of Applied Philosophy*, japp.12568. <https://doi.org/10.1111/japp.12568>

Ducato, R. (2020). Data protection, scientific research, and the role of information. *Computer Law & Security Review*, 37, 105412. <https://doi.org/10.1016/j.clsr.2020.105412>

Ernst & Young. (2019). *Realising the value of healthcare data: A framework for the future*. EY. [https://assets.ey.com/content/dam/ey-sites/ey-com/en\\_gl/topics/life-sciences/life-sciences-pdfs/ey-value-of-health-care-data-v20-final.pdf](https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/life-sciences/life-sciences-pdfs/ey-value-of-health-care-data-v20-final.pdf)

European Commission. (2021). *Commission Implementing Decision pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate protection of personal data by the United Kingdom*. [https://ec.europa.eu/info/sites/default/files/decision\\_on\\_the\\_adequate\\_protection\\_of\\_personal\\_data\\_by\\_the\\_united\\_kingdom\\_-\\_general\\_data\\_protection\\_regulation\\_en.pdf](https://ec.europa.eu/info/sites/default/files/decision_on_the_adequate_protection_of_personal_data_by_the_united_kingdom_-_general_data_protection_regulation_en.pdf)

European Data Protection Board. (2019). *Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of onlineservices to data subjects*. [https://edpb.europa.eu/sites/edpb/files/files/file1/edpb\\_guidelines-art\\_6-1-b-adopted\\_after\\_public\\_consultation\\_en.pdf](https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf)

European Data Protection Board. (2020). *Guidelines 05/2020 on consent under Regulation 2016/679*. [https://edpb.europa.eu/sites/edpb/files/files/file1/edpb\\_guidelines\\_202005\\_consent\\_en.pdf](https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf)

European Data Protection Supervisor. (2020). *A Preliminary Opinion on data protection and scientific research*. [https://edps.europa.eu/sites/edp/files/publication/20-01-06\\_opinion\\_research\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf)

Fairchild, A. L., Gable, L., Gostin, L. O., Bayer, R., Sweeney, P., & Janssen, R. S. (2007). Public Goods, Private Data: HIV and the History, Ethics, and Uses of Identifiable Public Health Information. *Public Health Reports*, 122(1\_suppl), 7–15. <https://doi.org/10.1177/003335490712205103>

Friis, R., Skov, L. R., Olsen, A., Appleton, K. M., Saulais, L., & Dinnella, C. (2017). Comparison of three nudge interventions (priming, default option, and perceived variety) to promote vegetable consumption in a self-service buffet setting. *PLoS ONE*, 12(5), 0176028. <https://doi.org/10.1371/journal.pone.0176028>

Georgieva, L., & Kuner, C. (2020). Article 9. Processing of special categories of personal data. In C. Kuner, L. A. Bygrave, & C. Docksey (Eds.), *The EU General Data Protection Regulation (GDPR): A Commentary* (pp. 365–384). Oxford University Press.

Hagman, W., Andersson, D., Västfjäll, D., & Tinghög, G. (2015). Public Views on Policies Involving Nudges. *Review of Philosophy and Psychology*, 6(3), 439–453. <https://doi.org/10.1007/s13164-015-0263-2>

- HiLo. (2021). *HiLo Maritime Risk Management*. <https://hilomrm.com/>
- Horn, R., & Kerasidou, A. (2020). Sharing whilst caring: Solidarity and public trust in a data-driven healthcare system. *BMC Medical Ethics*, 21(1), 110. <https://doi.org/10.1186/s12910-020-00553-8>
- Information Commissioner's Office. (2021). *Guide to the General Data Protection Regulation (GDPR)*. <https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-1.pdf>
- Kollnig, K., Binns, R., Van Kleek, M., Lyngs, U., Zhao, J., Tinsman, C., & Shadbolt, N. (2021). Before and after GDPR: Tracking in mobile apps. *Internet Policy Review*, 10(4). <https://doi.org/10.14763/2021.4.1611>
- Lades, L. K., & Delaney, L. (2022). Nudge FORGOOD. *Behavioural Public Policy*, 6(1), 75–94. <https://doi.org/10.1017/bpp.2019.53>
- McKee, S. (2014, October 8). *Care.data pilot schemes poised for launch*. PharmaTimes Online. [http://www.pharmatimes.com/news/care.data\\_pilot\\_schemes\\_poised\\_for\\_launch\\_1002595](http://www.pharmatimes.com/news/care.data_pilot_schemes_poised_for_launch_1002595)
- Merson, L., Phong, T. V., Nhan, L. N. T., Dung, N. T., Ngan, T. T. D., Kinh, N. V., Parker, M., & Bull, S. (2015). Trust, Respect, and Reciprocity: Informing Culturally Appropriate Data-Sharing Practice in Vietnam. *Journal of Empirical Research on Human Research Ethics*, 10(3), 251–263. <https://doi.org/10.1177/1556264615592387>
- Miller, J., Currie, S., & O'Carroll, R. E. (2019). 'If I donate my organs it's a gift, if you take them it's theft': A qualitative study of planned donor decisions under opt-out legislation. *BMC Public Health*, 19(1), 1463. <https://doi.org/10.1186/s12889-019-7774-1>
- National Academies of Sciences, Engineering, and Medicine. (2015). *Sharing Research Data to Improve Public Health in Africa: A Workshop Summary* (p. 21801). National Academies Press. <https://doi.org/10.17226/21801>
- National Data Guardian. (2021). *National Data Guardian statement on the General Practice Data for Planning and Research (GPDPR) programme*. Government of the United Kingdom. <https://www.gov.uk/government/news/national-data-guardian-statement-on-the-general-practice-data-for-planning-and-research-gpdpr-programme>
- N.H.S. Digital. (2021a, July 22). *GP Data for Planning and Research: Letter from Parliamentary Under Secretary of State for Health and Social Care to general practices in England*. <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/secretary-of-state-letter-to-general-practice>
- N.H.S. Digital. (2021b, August 24). *Collecting GP data—Advice for the public*. <https://web.archive.org/web/20220220064955/https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/advice-for-the-public>
- N.H.S. Digital. (2021c, August 24). *General Practice Data for Planning and Research: GP Practice Privacy Notice*. <https://web.archive.org/web/20211210163001/https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/gp-privacy-notice>
- N.H.S. Digital. (2021d, August 24). *General Practice Data for Planning and Research (GPDPR)*. <https://web.archive.org/web/20211210163001/https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/gp-privacy-notice>
- N.H.S. Digital. (2021e, August 24). *General Practice Data for Planning and Research: NHS Digital*

*Transparency Notice*. <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research/transparency-notice>

Nuffield Department Population Health. (2021). *RECOVERY: Randomised Evaluation for COVID-19 Therapy*. <https://www.recoverytrial.net/>

Office of National Statistics. (2021). *Accessing secure research data as an accredited researcher*. <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme>

O'Hara, K. (2019). *Data Trusts: Ethics, Architecture and Governance for Trustworthy Data Stewardship*. University of Southampton. <https://doi.org/10.5258/SOTON/WSI-WP001>

Open Banking. (2021). *Open Banking: The Future of Financial Services*. <https://www.openbanking.org.uk/>

Open Data Institute. (2021). *Data institutions*. Open Data Institute. <https://theodi.org/project/rd-data-institutions/>

OpenCorporates. (2021). *OpenCorporates: The largest open database of companies in the world*. <https://opencorporates.com/>

Oxford University. (2020). *Low-cost dexamethasone reduces death by up to one third in hospitalised patients with severe respiratory complications of COVID-19*. Oxford University News Release. [https://www.recoverytrial.net/files/recovery\\_dexamethasone\\_statement\\_160620\\_v2final.pdf](https://www.recoverytrial.net/files/recovery_dexamethasone_statement_160620_v2final.pdf)

Prabhu, P. K. (2019). Is presumed consent an ethically acceptable way of obtaining organs for transplant? *Journal of the Intensive Care Society*, 20(2), 92–97. <https://doi.org/10.1177/1751143718771711>

*Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (United Kingdom General Data Protection Regulation)*. (2016). European Parliament. <https://www.legislation.gov.uk/eur/2016/679/contents>

Saghai, Y. (2013). Salvaging the concept of nudge: Table 1. *Journal of Medical Ethics*, 39(8), 487–493. <https://doi.org/10.1136/medethics-2012-100727>

Scheibner, J., Raisaro, J. L., Troncoso-Pastoriza, J. R., Ienca, M., Fellay, J., Vayena, E., & Hubaux, J.-P. (2021). Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis. *Journal of Medical Internet Research*, 23(2), e25120. <https://doi.org/10.2196/25120>

Schmidt, A. T., & Engelen, B. (2020). The ethics of nudging: An overview. *Philosophy Compass*, 15(4). <https://doi.org/10.1111/phc3.12658>

Solon, J. (2014, July 2). A simple guide to care.data. *WIRED*. <https://www.wired.co.uk/article/a-simple-guide-to-care-data>

Starr, B. (2000). Organ procurement and mandated choice: An alternative to the existing system. *Princeton Journal of Bioethics*, 3, 70–80.

Sterckx, S., Rakic, V., Cockbain, J., & Borry, P. (2016). “You hoped we would sleep walk into accepting the collection of our data”: Controversies surrounding the UK care.data scheme and their wider relevance for biomedical research. *Medicine, Health Care and Philosophy*, 19(2), 177–190. <https://doi.org/10.1007/s11019-015-9661-6>

Temperton, J. (2016, July 6). NHS care.data scheme closed after years of controversy. *WIRED*. <http://www.wired.co.uk/article/care-data-nhs-england-closed>

Thaler, R. H. (2018). Nudge, not sludge. *Science*, 361(6401), 431–431. <https://doi.org/10.1126/science.aau9241>

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness* (Revised edition, new international edition). Penguin Books.

Triggle, N. (2018, May 24). 10 charts that show why the NHS is in trouble. *BBC News*. <https://www.bbc.co.uk/news/health-42572110>

U.K. Biobank. (2007). *UK Biobank Ethics and Governance Framework* (Report Version 3). <https://web.archive.org/web/20081227054532/http://www.ukbiobank.ac.uk/docs/EGF20082.pdf>

U.K. Biobank. (2021). *UK Biobank Homepage*. <https://www.ukbiobank.ac.uk/>

United Kingdom Parliament. (2018). *Data Protection Act 2018*. The National Archives. <https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted>

Verhenneman, G., Claes, K., Derèze, J. J., Herijgers, P., Mathieu, C., Rademakers, F. E., Reyda, R., & Vanautgaerden, M. (2020). How GDPR Enhances Transparency and Fosters Pseudonymisation in Academic Medical Research. *European Journal of Health Law*, 27(1), 35–57. <https://doi.org/10.1163/15718093-12251009>

Published by



ALEXANDER VON HUMBOLDT  
INSTITUTE FOR INTERNET  
AND SOCIETY

in cooperation with



CREATE



centre  
— internet  
— et  
— société



R&I

IN3

Internet  
interdisciplinary  
Institute

Universitat Oberta de Catalunya



UNIVERSITY OF TARTU  
Johan Skytte Institute of  
Political Studies