

An analysis of the social values involved in data linkage studies

Information governance in health

Research report

Sarah Clark and Albert Weale, Department of Political Science, School of Public Policy, University College London

August 2011

Acknowledgements

This report was written as part of a programme of work for an Economic and Social Research Council (ESRC) research project entitled 'Social Contract, Deliberative Democracy and Public Policy' (RES-051-27-0264). It was prepared in advance of a seminar held by the Nuffield Trust on 11 November 2010. The authors would like to thank all of those who attended the seminar for thought-provoking discussion, and to acknowledge the contribution made by the following people, all of whom offered very helpful comments on an earlier version of this report: Deryck Beyleveld, Roger Brownsword, Emma Byrne, Karen Dunnell, Trisha Greenhalgh, Joan Higgins, Søren Holm, Graeme Laurie, Onora O'Neill and Nayha Sethi. Naturally, any remaining mistakes are our own.

Contents

Executive summary	4
1. Introduction	5
2. Current law and policy	7
3. Social values and information governance	15
4. Conclusion	28
Bibliography and references	31
Appendix 1: Further details and discussion of legal instruments	33
Appendix 2: Glossary of terms	37
Appendix 3: International comparisons of data protection	39

Executive summary

1. Medical and health services research is now often carried out by linking patient-level data, relying on information that contains individual identifiers. Within the dominant ‘consent or anonymise’ interpretation of the legal framework, if data contain any information that might identify a data subject – that is, if the data are not fully anonymised – explicit consent must be sought.
2. There are various provisions in law for use of data without explicit consent, one of which is a ‘public interest’ justification. However, the terms of these justifications lack clarity and researchers have been reluctant to utilise them as a result, opting instead to exercise caution and seek explicit consent in order to avoid the possibility of legal sanctions.
3. Seeking consent in this way can impose burdens on researchers which, some have argued, obstruct the success of research. However, others urge the importance of maintaining the emphasis on consent as a way to ensure privacy.
4. In the current governance framework, consent can be thought of as a form of ‘counter-control’ by which the obligations of confidentiality of data controllers such as general practitioners (GPs) and hospital staff are waived. Rights to informational privacy are thus conceptualised as rights of control over information about ourselves.
5. A number of social values are advanced by this model. It recognises the value of privacy for individuals and seeks to grant them autonomy over what happens to their personal information. Protecting privacy and granting autonomy may be seen as important aspects of what we owe to people if we are to treat them with respect.
6. The public benefits of medical and health services research may provide a justification for weakening the consent as counter-control principle: citizens accept some limitations on their individual privacy in other areas of life in order to obtain certain collective benefits.
7. Balancing individual rights and collective benefits in this way involves conceptions of fairness and reciprocity, the terms of which may depend on the quality and distribution of benefits and on the existence of institutional contexts in which certain terms of cooperation are already implied.

1. Introduction

Research using patient data holds the promise of improving human health in many forms, via epidemiological studies, public health surveillance, monitoring of drug safety, improved health service management and evaluation of surgical interventions. In Box 1, we give some examples. Such research is often carried out using databases, so that analysts can examine large quantities of data quickly, at comparatively low cost, and without any interference in the lives of data subjects. However, such research often requires data linkage, that is to say, matching and combining data from multiple databases. Such data linkage cannot be done with fully anonymised data, but requires some form of individual identifier to enable matching.

According to the dominant ‘consent or anonymise’ interpretation of the legal framework in the United Kingdom (UK), in particular of the Data Protection Act 1998 (DPA), if data contain any information that might identify a data subject – that is, if the data are not anonymised – it is thought that explicit consent must be sought from that individual (see Al-Shahi, 2000; Academy of Medical Sciences, 2006; Haynes and others, 2007). Seeking consent in this way can raise a series of problems for researchers: it can be expensive in terms of both financial and time resources, it can cause delays in starting research work, and it can jeopardise the validity of the outcomes of research due to factors such as consent bias and incomplete samples. Conversely, full anonymisation of data renders the data useless for much research, including any that relies on techniques of data linkage.

There are, then, considerable *practical* problems with a polarised ‘consent or anonymise’ approach, and some argue that it prevents useful research being conducted (see, for example, Strobl and others, 2000; Peto and others, 2004; Academy of Medical Sciences, 2006; Haynes and others, 2007). However, others concerned with threats to personal privacy urge the continued importance of informed and explicit consent in relation to the processing and use of patient-identifiable data in cases where full anonymisation is impossible (see, for example, Health Select Committee, 2007, EV 103 and 106; and the recommendations of the European Commission-funded PRIVIREAL Project¹).

‘Consent or anonymise’ is, however, one *interpretation* of the legal framework rather than a strict legal requirement. While the DPA and the wider regulatory regime seek to protect individual privacy and confidentiality by stressing the need for patient consent to certain types of data usage, they also recognise and protect the public interest in research by allowing some forms of research to be undertaken using identifiable data but without requiring explicit consent – indeed, much health services research in the NHS is already carried out under these provisions. The Information Commissioner has acknowledged that ‘consent or anonymise’ is an interpretation of law rather than

1. The PRIVIREAL Project is a European Commission Framework 5 funded project examining the implementation of the European Union (EU) Data Protection Directive 95/46/EC in relation to medical research and the role of ethics committees. It has recommended a broader definition of personal data to encompass forms of anonymised data; a stricter interpretation of the exemptions under which medical research might be carried out; and tighter checks when researchers use anonymised data or rely on exemptions. Recommendations are available online at www.privireal.org/content/recommendations/

a requirement, suggesting that “it is a common misconception that the Act [the DPA] always requires the consent of data subjects to the processing of their data” (Information Commissioner’s Office, 2002). There are alternatives to ‘consent or anonymise’ within the regulatory regime but, it has been suggested, a conservative interpretation of the law has been adopted by researchers and health service bodies out of confusion over legal provisions and professional guidance, and out of fear of litigation (Information Commissioner’s Office, 2002).

At the heart of this debate is a discussion about how best to realise certain social values, including individuals’ interest in privacy and autonomy, on the one hand, and their interests in obtaining the benefits of research and improved health service planning, on the other. The purpose of this report is to review these social values as they apply to questions of information governance, using the methods of normative policy analysis. It is part of a larger project to explore how a ‘social contract’ might be written that strikes a fair balance between individual rights and obligations, as well as securing public acceptability. It also links to wider debates in public policy about the role of social values in decision-making, for example the National Institute for Health and Clinical Excellence’s statement of *Social Value Judgements* (NICE, 2008) and the forthcoming analysis by the Charity Commission of the public benefit test (Charity Commission, 2011).

Much of the existing discussion of data protection in health care has been focused on genetic information (for example, Human Genetics Commission, 2002). There is a large debate in bioethics about ‘genetic exceptionalism’, and the extent to which genetic information is distinct. Without intruding into that debate, we focus on data linkage that by and large does not involve the use of genetic information, but rather looks at, for example, data on diseases, effectiveness of interventions or patterns of service use.

Box 1: Examples of research using data linkage

- In a study in Scotland, researchers linked national pregnancy and perinatal death registries to undertake a retrospective cohort study of 130,000 women having a second birth whose first birth had involved complications. The study indicated that women who experienced pre-term birth, delivery of a low-weight baby or preeclampsia in a first birth were also at increased risk of unexplained stillbirth in the next pregnancy. As a result, mothers-to-be who are at risk are now identified and offered early interventions to pre-empt problems in their second pregnancy (Smith and others, 2007).
- The Bristol Royal Infirmary Inquiry found that there were inadequate procedures in place to ensure the systematic follow-up of patients following paediatric cardiac surgery. In particular, it was thought that Hospital Episode Statistics (HES) data on deaths within 30 days of surgery were unreliable and could substantially underestimate post-operative mortality. HES data were linked to the Office for National Statistics’ register of deaths to evaluate the reliability with which HES captured post-operative mortality and to estimate the number of deaths occurring after discharge from hospital (Murray and others, 2000).
- HES, GP and Accident & Emergency data have been linked to produce predictive models that can identify individual patients at high risk of unplanned hospitalisation or admission into social care. The models use pseudonymised data but that data are re-identified in order for the relevant patients to be contacted by their GP and offered appropriate interventions. Identifying these individuals allows early interventions to be offered that can prevent hospitalisation and promote independent living for longer (Lewis and others, 2008).

2. Current law and policy

This chapter:

- sets out the legal framework that regulates data use, focusing primarily on the Data Protection Act (DPA) but also outlining other legal provisions (full details of the DPA and further discussion of the other most relevant statutory instruments can be found in Appendix 1)
- identifies and discusses the available legitimations for use of ‘identifiable’ data, questioning the presumptions of the ‘consent or anonymise’ approach
- considers issues around anonymisation of data.

A glossary of terms used in this section can be found in Appendix 2. Additionally, examples of data protection regimes in other countries can be found in Appendix 3.

Introduction

Until as recently as the 1990s, there was no freestanding legal right to privacy, and professional guidelines encouraged researchers to view research using patient information as a professional ethical duty, placing little emphasis on seeking consent or the approval of research ethics committees, so long as no harm was done to the patients in question (see, for example, Royal College of Physicians Committee on Ethical Issues in Medicine, 1994). However, in the years following the introduction of the DPA and the Human Rights Act 1998, and with health policy increasingly advocating a ‘patient-focused’ health service (see, for example, DH, 2004), the interests of patients have increasingly been understood in terms of individual rights.

In practice, this has led to an interpretation of data protection as ‘consent or anonymise’, in which individual consent is thought to be the only appropriate way of justifying the use of identifiable data – that is, data that are not fully anonymised – in research. Fully anonymised data are not useful in any research that employs techniques of data linkage, owing to the need for identifiers to enable matching and statistical verification. Instead, pseudonymised data are often used: these are data that have had all personal identifiers (information such as name, address and so on) removed but have been allocated a code number that enables the data controller to link the data back to the individual via a ‘key’ which decodes the data. However, this type of data is still classified as ‘identifiable’.

Under the ‘consent or anonymise’ regime, which it has been suggested has come to dominate governance of health information (Academy of Medical Sciences, 2006), researchers wishing to use identifiable data require explicit consent from data subjects or, where that is impossible, they must obtain permission to do so via special exemptions granted by application to the National Information Governance Board (NIGB). Either of these routes can involve lengthy processes, significant additions to cost and delays to research projects; the prospect of which may deter research in the first place (see, for example, Haynes and others, 2007).

Legal framework

The principal legal provisions that regulate information governance in the UK are as follows:

- the DPA 1998
- the Human Rights Act 1998
- the Freedom of Information Act 2000
- the National Health Service Act 2006 (previously Section 60 of the Health and Social Care Act 2001)
- the common law duty of confidentiality.

Professional guidelines that interpret the legal framework have been produced by the Medical Research Council (MRC, 2000), the Department of Health (DH, 2003), the British Medical Association (BMA, 2007) and the General Medical Council (GMC, 2009). Additionally, in the NHS, a system of Caldicott Guardians was established by the 1997 Caldicott Report (DH, 1997). Caldicott Guardians are responsible for safeguarding the confidentiality of patient information within the NHS and for overseeing information sharing.

The aims of the legal framework overall are, broadly:

- to regulate data flows
- to protect a patient's right to privacy
- to enforce duties of confidentiality
- to protect the public interest in the benefits of research.

The Data Protection Act

The DPA establishes a series of principles by which personal data must be processed, maintained and transferred if use of such data is to be considered fair, lawful and proportionate (these principles and the relevant additional schedules are summarised in Appendix 1). The original purpose of the European Union (EU) Data Protection Directive from which the DPA is derived, was to facilitate data flows while also protecting individual privacy.

The DPA is primarily concerned with 'personal data', defined as "data which relates to a living individual who can be identified – a) from those data, or b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller" (DPA, Section 1 (1)).

Information about physical or mental health is considered by the Act to be both 'personal' and 'sensitive'. Health data employed in data linkage research will therefore fall into both of these categories, since it concerns information about physical and mental health, and it is not fully anonymised but rather pseudonymised, containing individual identifiers to enable matching.

The fair and lawful processing of sensitive personal data requires compliance with the eight general principles of the DPA (see Appendix 1). Most notable for our purposes are the first two principles.

The first principle requires that processing of data must meet at least one condition for ‘fair and lawful processing’ set out in each of Schedules 2 and 3 of the Act, although, in general, any research that meets a condition of Schedule 3 will also meet at least one requirement of Schedule 2.

To summarise, Schedule 2 of the Act permits processing of information under four main conditions:

1. where consent has been obtained
2. if processing is in the vital interests of the data subject
3. if processing is “necessary for the exercise of... functions of a public nature exercised in the public interest by any person”
4. if the processing is necessary for the purposes of legitimate interests pursued by the data controller or by the third party or parties to whom the data are disclosed.

Schedule 3 of the Act provides three main conditions under which sensitive personal data can be fairly and lawfully processed:

1. where explicit consent has been obtained
2. if processing is necessary for protecting the vital interests of the data subject or another person *and* it is either impossible or unreasonable for explicit consent to be obtained
3. if processing is necessary for medical purposes and is undertaken by a medical professional, or equivalent other professional who owes a duty of confidentiality with regard to patient information.

The second principle of the DPA specifies that “[p]ersonal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes”. Hence, with health information originally obtained for the purposes of medical treatment, subsequent use for the purposes of research can mean that either consent or another form of legitimation for use is obtained by researchers. It might be argued that when patients provide information for the purpose of their own treatment, it is then ‘incompatible’ to use the data for the purposes of research or public health monitoring. However, the situation is additionally complicated by Section 33 (2) of the Act (see below) in relation to statistical and historical purposes.

Other relevant legal and policy provisions

Two other related legal provisions are particularly important alongside the DPA.

The first is the common law duty of confidentiality, where information that is disclosed in a situation of confidentiality – most obviously the doctor–patient relationship – should not be disclosed to third parties. This duty can be overridden in some cases, for example the notification of certain diseases, but also, since perfectly innocent personal information may be passed from one professional to another, for example, about possible complications, it is accepted that breaches of confidence are only actionable where they are “an abuse or unconscionable to a reasonable man” (Home Office, 2003).

The second is Section 251 of the National Health Service Act 2006 (previously Section 60 of the Health and Social Care Act 2001), which modifies the common law with the explicit intention of facilitating research. Section 251 provides a power to allow patient-identifiable information needed for ‘medical purposes’ (defined as including medical research and the management of health and social care services) to be used without the consent of patients. The power can be used only where obtaining consent is not feasible and where anonymised information will not suffice.² Lord Falconer of Thornton (2001), the Cabinet minister responsible for the introduction of the original Section 60 provision, commented that it offered “a power to ensure that personal information needed to support essential research activity can be provided without the consent of individuals and without breaching the common law of confidentiality”.

Current guidance on confidentiality issued by the BMA, the GMC and the MRC appears to be relatively harmonious with this and other legal provisions governing the use of patient information. The most recent GMC guidance on disclosing patient information in effect lists which disclosures would not put a doctor in breach of the common law duty of confidentiality, and it reads as follows:

For many secondary uses, it will be sufficient and practicable to disclose only anonymised or coded information. When identifiable information is needed, or it is not practicable to remove identifiable information, it will often be perfectly practicable to get patients’ express consent.

You may disclose identifiable information without consent if it is required by law, if it is approved under Section 251 of the NHS Act 2006, or if it can be justified in the public interest and it is either:

- a. necessary to use identifiable information, or*
- b. not practicable to anonymise or code the information*

and in either case, not practicable to seek consent (or efforts to seek consent have been unsuccessful).

(GMC, 2009, p 18.)

2. This provision, first introduced in the form of Section 60 of the Health and Social Care Act 2001, was a response to guidelines on confidentiality published by the GMC in 2000, which had effectively prohibited any data transfers, even those to cancer registries, without the patient’s express consent. This threatened to jeopardise reporting of cancer incidences as the guidance implied that doctors who made such reports, in what were by then well-established procedures, would now face disciplinary proceedings. Data collection arrangements predictably began to fail not only for cancer reporting but also for communicable diseases, and many research projects were delayed or blocked.

Possible legitimations for the use of identifiable data

This section brings together the various legitimations for use of data that are available from across the legal framework. These are: consent, processing data for ‘medical purposes’, processing data in the public interest, and use of data for ‘statistical or historical purposes’. It also highlights some issues arising in relation to each of these possible legitimations.

Consent

Consent is not defined in the DPA, but the implication is that where consent is the only justification a researcher has for processing health data, that consent must be explicit and informed. Yet for researchers, seeking explicit consent can be highly problematic for a number of reasons. Rarely can explicit consent be obtained at the point when information is originally taken from a patient because the future possibilities for research using that information cannot be anticipated in advance, so it is often necessary to contact data subjects sometimes many years later to request consent. Given that large-scale data linkage studies can involve data from hundreds or even thousands of data subjects, re-contacting each individual can be simply impossible (for example, contact details may no longer be valid or subjects may no longer be alive), prohibitively expensive or administratively impractical. Problems of selection bias may also arise, when some groups of people within a sample are more likely to refuse to give consent than others, thereby skewing the entire sample and giving rise to problems of validity in statistical inference (see, for example, Cox and others, 2001). Thus, requiring explicit consent may mean that the validity of research is compromised to a degree whereby its quality is significantly reduced, and the benefits it can therefore provide to patients and society at large are also reduced.

Processing data for ‘medical purposes’

However, explicit consent is only *one* of the means to justify use of sensitive personal data. The ‘medical purposes’ legitimation outlined in Schedule 3 of the DPA is especially significant, but also open to interpretation. “Medical purposes”, as defined in the Act, include “preventive medicine, medical diagnosis, medical research, the provision of care and treatment and the management of health care services” (DPA, Schedule 3, 8 (1) and 8 (2)). The explicit inclusion of ‘medical research’ here is notable. However, nowhere in the Act is the term ‘medical research’ defined. Similarly, what the term ‘necessary’ means in the phrase ‘processing is necessary for medical purposes’ is also left unspecified. It seems likely that data linkage using identifiable information that is carried out in the context of NHS health services research will fall within the scope of ‘medical purposes’, and that it could be conceived as ‘necessary’ in so far as it is required for the effective functioning of the health service. It has been suggested that invoicing, accounting and statistics are elements of the management of health care services, but that processing of insurance claims are not (see European Commission Data Protection Working Party, 2007: pp10–12).

According to the opinion of an independent EU advisory body on data protection, ‘medical purposes’ legitimation only covers activities that are required for the direct provision of services and not other research, even in the areas of public health and social protection (although these may be covered by Article 8 (4) on public interest exemptions). Further, that opinion states that “the processing of personal data on grounds of Article 8 (3) must be *required* for the purposes of provision of services, and not merely ‘useful’” (Article 29 Data Protection Working Party, 2007: pp10–12, emphasis in original). However, it has also been argued that any medical research

that has been approved by a Research Ethics Committee, whether required for the provision of services or not, must be, by definition, *necessary* since it is simply not ethically appropriate to conduct *unnecessary* research (Academy of Medical Sciences, 2006: p25). Therefore, there is a sense in which there is an unnecessary ‘doubling up’ of legitimization in data use.

The other potentially problematic aspect of Schedule 3 is that which relates to the processing of data only by a ‘health professional’ or ‘a person who owes an equivalent duty of confidentiality’. This appears to be a problem where researchers are not also health professionals. The term ‘health professional’ includes registered health professionals, such as doctors, nurses and therapists, but also senior scientists within health service bodies. However, it has been suggested that in fact almost all researchers employed by institutions are under obligations of confidentiality because of the legal principles within the common law of confidentiality. Research contracts typically include clauses that make misuse of confidential information a cause for disciplinary proceedings to reflect these obligations. As such, it has been argued, all researchers should be able to conduct research under the ‘processing by a health professional’ exemption (see Academy of Medical Sciences, 2006: p25). Once again, there is a possible ‘doubling’ of requirements here.

Processing data in the public interest

Another justification for research using identifiable data without explicit consent is that offered by the ‘public interest’ exemption provided by Schedule 2 of the DPA. This allows derogation from the requirements around processing sensitive data, so long as suitable safeguards are employed and the reasons for research are those of the public interest. What constitutes ‘public interest’ is not defined, however. The provision for derogation is expanded on in guidance to the EU Data Protection Directive, which states that “Member States must also be authorized, when justified by grounds of important public interest, to derogate from the prohibition on processing sensitive categories of data where important reasons of public interest so justify in areas such as public health and social protection... scientific research and government statistics” (EU Data Protection Directive 95/46/EC, Recital 34). Once again, however, it seems plausible to suggest that any research that has been approved by a properly constituted Research Ethics Committee will, perforce, be in the public interest – if it were not, then it would not be ethically appropriate and would not receive approval.

Use of data for ‘statistical or historical purposes’

One further alternative to consent as a legitimization for processing sensitive data is provided by Section 33 of the DPA. This exempts data processors from a number of obligations where research is for ‘statistical or historical purposes’. It demands that researchers provide information to data subjects if they request it, but not that they must do so as a prerequisite for conducting the activities, and that data is only processed for lawful purposes and not in order to support decision-making with respect to particular individuals or in a way that causes damage or distress to them (DPA 1998, Section 33). Importantly, if the purpose of the research processing meets these requirements, it is exempt from the second data protection principle, meaning that personal data can be processed for purposes other than for which they were originally obtained.

Data linkage studies could plausibly be seen to fit into this category: the data in these studies are often being processed for statistical purposes and, because of those

purposes (and presuming adequate data security procedures), are highly unlikely to cause damage or distress to individuals, and are not used to make clinical decisions about them.³

The issue of anonymisation

Another source of confusion in the DPA concerns the issue of anonymisation. The Act does not apply to anonymised data: the EU Data Protection Directive clearly states that “the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable” (EU Data Protection Directive 95/46/EC, Recital 26). There is considered to be negligible interference in the patient’s privacy and no breach of confidentiality on the part of doctors who release anonymised information for research purposes.⁴ Anonymised data have, in effect, been stripped of any elements that would allow identification of individual patients and contain no means by which to re-identify data.

However, neither the EU Directive nor the DPA defines what is to be classed as anonymised data, or explains how data are to be anonymised, given that whoever performs the process will necessarily have access to identifying information in order to de-identify it. Guidance from the Information Commissioner’s Office (2002) states that, although anonymous data may fall outside the remit of the DPA, the *act* of anonymisation does not: “In anonymising personal data the data controller will be processing such data and, in respect of such processing, will still need to comply with the provisions of the Act”. So it seems that a rather peculiar situation exists where, in order to anonymise data so that consent of data subjects is not required, consent of those data subjects is required.

However, it is possible that researchers may use (and anonymise) data without consent of the data subjects if they can comply with one of the other Schedule 3 conditions, most likely that of processing data for medical purposes if such processing is carried out by a health professional or a person who owes an equivalent duty of confidentiality.

It has been argued by the Confidentiality and Security Advisory Group for Scotland (2002: p32) that: “It is important to note that while there are no legal restrictions on the use of data that do not identify patients, they do have a right to know when it is intended that their information will be anonymised for a range of appropriate purposes”. This concern reflects some views expressed in a recent consultation on public perceptions of secondary data use (DH, 2010). Only one in six of those who responded to the consultation (around 700 responses were received) felt that it was acceptable for anonymised data from ‘sealed envelopes’ – sections of the Electronic Patient Record to which patients will be able to restrict access – to be used for research purposes without consent, even with all appropriate safeguards in place. Some individuals and groups who responded, especially those with ‘sensitive’ conditions, were strongly opposed to use of their data even in anonymised form. One such respondent suggested that anonymised data should not be used without consent because “[w]e all think we should know when people look through our records because it intrudes on our privacy” (DH, 2010: p26). However, there are other studies that give a different picture and there is a general question about how to aggregate differing

3. An interesting case here is that of predictive risk analyses, which use large datasets but with the aim of identifying individuals who are at high risk, for example, of admission to hospital or social care. In this case, Section 33 exemptions would seem inapplicable.

4. It is feasible that under the definition of ‘medical purposes’ in Schedule 3 of the Act, doctors may be able to release identifiable information for the purposes of health services research, but only if it can be guaranteed that such information will be processed only by professionals who share a duty of confidentiality to the patient.

opinions as well as problems of statistical inference to the population at large from unrepresentative samples (see, for example, Barrett and others, 2006).

Summary

- The general way in which the DPA has been interpreted is in terms of a ‘consent or anonymise’ approach, but this is only one approach, although it has the weight of practice.
- There are difficulties for researchers in securing adequate explicit consent for use of data that are not fully anonymised.
- But, contrary to the impression given by the ‘consent or anonymise’ interpretation, alternative legitimations for processing data do exist. These include processing of data for ‘medical purposes’, processing data in the public interest, and use of data for ‘statistical or historical purposes’. These provisions in the legal framework demonstrate recognition that there is a legitimate interest in research.
- However, the terms of these alternative justifications for use of data are inadequately defined, and this leaves researchers unsure as to whether their use of data can be legitimated in these ways.
- While the ‘consent or anonymise’ approach may present anonymisation as the primary alternative to consent, the process of anonymisation is not legally straightforward.

3. Social values and information governance

This chapter analyses the social values involved in questions of information governance in health. These values are, as we see them:

- consent
- privacy
- autonomy
- property rights
- the duty of confidentiality
- public benefit
- fairness and reciprocity.

Introduction

In the previous chapter, we noted that ‘consent or anonymise’ is only one interpretation of the law. In this chapter, we examine the social values that the law may serve and see how far this interpretation meets those values.

In general, the relevant social values involve a concern for the promotion of the common good combined with an insistence on the idea of respect for persons. As citizens and individuals, we want our data to be kept safe, but we also wish for the benefits of research, and sharing data under appropriate conditions is a way of securing this. We fail to respect people not only when their interests are damaged or when they are exposed to insecurities that leave them open to harm. Rather, what is involved in unjustified data sharing is a lack of respect for the person who is the data subject. It is like gossiping behind someone’s back: it shows a level of discourtesy which amounts to a lack of respect. Of course, merely conveying to one person information about a third party is not in itself a manifestation of a lack of respect. But when that information contains the sort of details it is reasonable to assume that someone would rather be kept confidential, or which are shared in a specific context of confidentiality, then a lack of respect is shown. In the rest of this chapter we examine the arguments that link respect in this sense with our interests in the benefits of research.

The seven values that we identify as relevant are consent, privacy, autonomy, property rights, the duty of confidentiality, public benefit, and fairness and reciprocity.

Consent

The mechanism by which patient control is exercised in the ‘consent or anonymise’ view of informational privacy, is explicit, informed consent – that is, consent informed by an understanding of what it is that is being consented to. The Data Protection Act (DPA) holds that the use of any personal information, defined as information or links to information that make an individual ‘identifiable’, must be consented to by the data subject, in the absence of other justifications such as those related to the public interest. Consent thus acts as a form of waiver whereby the data subject agrees to the use of information, which use, it is implied, would otherwise be impermissible.

However, there is no absolute ‘right to consent’ or ‘right to withhold consent’ in the same way that there are moral and legal rights to privacy or confidentiality. Rather, consent is a sign from an individual that they are willing to accept an action that would otherwise interfere with their moral or legal rights – such as those of privacy or confidentiality. That is, consent is a justification or authorisation for action, rather than an absolute right or requirement. One obvious implication of this is that not *all* actions require consent, only those actions where justification is needed. So consent is unnecessary where no moral or legal right is at stake, such as in circumstances where anonymised data are used in research and the impact on individual privacy is negligible. Further, considered in this way, it becomes clearer to see that because consent is not an *absolute* right or requirement, it may be outweighed by other justifications for what may otherwise be viewed as an interference in privacy. That is to say that consent is but *one* justification for such actions.

Nonetheless, explicit consent appears to put the data subject in control, in as much as they give their authorisation for use of information. It is not clear that the kind of control over data that is implied by the dominant model is possible, however, given the nature of consent and its limitations. One reason for this is the difficulty in specifying precisely and determinately what it is for which consent is being sought, and ensuring that patients’ understandings of what is being asked for are fully consistent with what those who are requesting their consent understand it to be (see O’Neill, 2003). For example, a research programme may propose to investigate the links between a certain surgical intervention and subsequent rates of hospital admissions, and describe that proposal in some detail for the purposes of patients’ consent. Certain aspects of the proposal may implicitly require other procedures not specified in the proposal, because they are taken for granted by researchers as precisely an implicit part of that proposal. However, for the patient who is being asked to consent to the use of their data, that such procedures may take place will not be evident – patients rarely share the same level of medical understanding as their clinicians.

A further complication is that researchers themselves are rarely the people to obtain consent from a data subject. Usually, consent will be sought by a clinician treating the patient and information will then be uploaded on to databases or passed on in other appropriate forms to researchers. Two problems arise from this. First, the researchers cannot themselves be certain that the consent obtained was truly informed or that the person had the capacity to provide it. Second, the *clinical* context in which consent is obtained may carry certain meanings for the patient that are not appropriately applied in consenting to *non-clinical* uses of their information: patients expect health professionals to be acting in *their* best interests and to be motivated by improving *their* health rather than the health of the population or the achievement of advances in medical knowledge.

Consent given by patients may be affected by this presumption and – out of trust for their clinician – lead to them either not paying due attention to what they are being asked to consent to or mistakenly consenting on the basis that some therapeutic benefit may come to them directly as a result. O'Neill (2003) has described many problems with consent as arising from it being a 'propositional attitude' where an individual can agree to general proposals but where it is impossible to specify exact limits to what has been agreed to and where, consequently, it cannot be assumed that the full range of the individual's interests are protected. Once again, to view consent as a form of control over information seems to ignore the conditions under which it is often given, where a patient may not be fully cognisant of what is being proposed to them, for reasons associated with the context in which consent is sought. Courts have been sceptical about the quality and genuineness of consent when it is not obtained by researchers themselves, so given that it can rarely be so obtained, researchers are often left in a no-win situation (Academy of Medical Sciences, 2006).

Privacy

One obvious value that should be served by any data protection system is protecting the interest in privacy that individuals legitimately have in their own lives. Just as we wish to have privacy in our physical space, so we can protect our psychological self by informational privacy. What is valuable in privacy is, as Laurie (2002: p8) puts it, a sense of separateness from others and 'being apart from', a state of non-access to an individual's physical or psychological self.

Informational privacy in health care obviously cannot be absolute. Data need to be shared with medical teams, so it is not privacy as such that is at issue, but rather the risk of information being shared with others who are not under a duty to behave in our best interests. Sometimes this concern rests on there being some facts about us that we do not wish others to know because it may cause us social embarrassment or because we feel that such information is simply none of their business. Such psychological unease is plausibly linked to fears about being judged by others and how we might be the subject of their prejudice, but it may also be linked to fears about more tangible harms: it is possible that the release of information about one's lifestyle, health, finances or habits could result in harm to one's reputation, to discrimination or to detrimental effects on one's employment prospects, if the relevant information is sufficiently widely circulated. Our personal relationships too can be affected by certain information about us being made known to our nearest and dearest when we might otherwise have sought to protect them from that knowledge, or at least to share it with them in a manner of our own choosing.

As we have seen, the DPA defines personal information as that which makes a person 'identifiable'. Many facts about me may make me identifiable, but surely not all of them would count as 'personal' information. There may be numerous facts that are already available and in the public domain: perhaps my strongest identifier, my name, for instance, will appear on an electoral register accessible to political parties or on the databases of any number of utility companies. My identity might easily come to be known by a person who happens to see me at my place of work, or entering my house. Can such information really be termed 'personal', with the implication that I should have to give consent for others to know it? That would seem at least impractical and, more likely, quite absurd. Perhaps what is troubling about identifiability is that it gives some *unique* information about a person, although this too seems unsatisfactory:

as Manson and O'Neill (2007: p103) observe, there may be lots of people who have cancer, but each of them may well consider this fact about them to be a private, 'personal' matter.

A common perception of what constitutes 'personal' information is that it is 'sensitive' or intimate in a way that might make it particularly socially embarrassing or damaging if it were known by others or used to our disadvantage: such information might include facts about our sexuality, religious beliefs, financial status and health status. In terms of the vulnerability that informational privacy rights might be thought to protect, it is perhaps with regard to these issues that individuals feel most exposed. However, what is 'sensitive' and 'intimate' to one person may be mundane to another and vice versa, so determining a comprehensive list would be difficult. However, that such variability of feeling exists about these issues is perhaps in itself grounds for imposing consent requirements in order to allow for individuals weighing the value of such information differently: some people may feel it to be sensitive, others may not, but for those who do, it will likely be important that it is treated accordingly. Not to do so, or at least not to offer the opportunity to do so, could be to risk promoting a sense of intrusion into areas of intimate life, which in turn could jeopardise trust in the relevant public policy.

A further problem arises from defining personal information as 'sensitive' or 'intimate'. Some kinds of information commonly thought to be in these categories – for example, facts about sexual or mental health status – are ones that are often felt to be the subject of social stigma. While, on the one hand, giving individuals special rights of control over the use and communication of such information may be viewed as a means of protecting them from potential stigma, on the other, it could serve to further entrench that stigma, both by implying that there is indeed something to be ashamed about in relation to such issues and by failing to raise any challenge to the status quo, which maintains certain shameful and shaming associations. There are practical harms that might occur here, for example, in the shape of insurance or employment disadvantages, but also harms that are such because they involve some kind of affront to autonomy – the harm arises out of something being done, which a person does not wish to be done.

Research into public perceptions of the use of 'sensitive' data shows mixed attitudes. A recent Department of Health consultation enquired about the use of data contained in 'sealed envelopes': these are proposed sections of the Electronic Patient Record where patients can opt to store information to which they wish access to be restricted or barred. It is likely that the information patients choose to store in this way will be felt by them to be 'sensitive' in some way, perhaps pertaining to sexual or mental health issues. The consultation showed that some of those representing relevant patient groups, such as the Royal College of Psychiatrists and the Terrence Higgins Trust, believed that such information should never be used, even in anonymised form, without the patient's explicit consent: the Terrence Higgins Trust stated that "data which has been placed in a 'sealed envelope' should not be accessed for research purposes, even in anonymised form" (DH, 2010: p22). However, others recognised that withholding information in this way could bias research and thereby hinder greater understanding of the relevant conditions: the Medicines and Healthcare products Regulatory Agency (MHRA) commented that "[f]ailure to have access to this data may provide false results which is potentially much worse than not having access to data at all" (DH, 2010: p21).

Describing information as ‘personal’ perhaps also obscures the likely nature of the use of that information, for example, in research. Our fears of others using facts about us to our disadvantage might be relevant if our information were being handed out to local gossips with an interest in maligning reputations: then the use of our information would indeed be ‘personal’ and may come to adversely affect us *personally*. However, in the contexts of medical and health services research, that certain information is about any given individual is of no interest to researchers, and they have nothing to gain by using information about that individual to their disadvantage. Research does not aim to find out about particular individuals but rather to use data about many individuals in order to draw conclusions that might be applicable to large populations. When the ends of research have no bearing on us as individuals and where there is no intention to identify us personally, the identification of information as ‘personal’ seems, in this dimension as in others, misleading.

Autonomy

Some link the idea of respect for persons to the idea of autonomy, understood as the claim that an important element in human flourishing is self-definition, or at least having control over a significant portion of one’s life. According to Benn (1971: p239), for example, “respect for someone as a person, as a chooser, implies respect for him as one engaged on a kind of self-creative enterprise, which could be disrupted, distorted, or frustrated even by so limited an intrusion as watching” (see also Dworkin, 1993). A concern for autonomy in this sense is not co-extensive with concern for privacy, since there are grounds for respecting privacy even when the conditions of autonomy do not prevail. Thus, there is just as much a problem of values when the privacy of comatose patients is breached as when the privacy of those who are fully conscious is breached, perhaps even more so since there is no opportunity to consult the comatose about their wishes.

Nonetheless, that there is a link between privacy, respect for persons and autonomy is not hard to see. The Kantian injunction to treat persons not merely as means but also as ends in themselves is not far below the surface here: using information about a person without their consent for one’s own purposes could be to treat them merely as means and therefore fail to show them the respect they are due as an agent who is an end in themselves. Consent, arguably, functions as a vehicle for the exercise of autonomy and so (also arguably) indicates a respect for persons that allows information to be used without treating those persons *merely* as means to an end. So, if privacy is a zone of separateness, respect for which is a component of respect for persons, then that zone should be put under the control of persons and only incurred upon with their consent.

Hence, rights to informational privacy are commonly viewed as rights of control: that is, rights to control *what* information we share with *which* others, and rights against those others that they use the information we share with them only in specific ways. Such rights seem to offer a means of protection against vulnerability and of minimising the threat of harm that could be done by another using our information to our disadvantage. They do so by allowing us to restrict who has access to which information about us and what they may legitimately do with that information. We think that it is this idea of autonomy that lies behind the ‘consent or anonymise’ model.

Property rights

It is perhaps more than mere coincidence that the conception of informational privacy rights as a form of control seems to have developed alongside an increasing emphasis on patients' rights, consent and involvement in treatment in clinical settings, and alongside the development of freer access to information via the internet. A recent report from the Nuffield Council on Bioethics (2010), *Medical Profiling and Online Medicine*, has shown the variety of uses of information technology from online health information, online personal health records, online purchasing of pharmaceuticals to personal genetic profiling for disease susceptibility. These developments are sometimes linked to the idea of 'personalised medicine' although the Nuffield report shows that this is too simple an idea. However, there is no doubt that there are significant trends in contemporary culture that interpret information as a form of property.

However, there are a number of confusions that are possible here: while the notion of control is not synonymous with that of ownership, the latter could plausibly be taken to be implied by the former, albeit wrongly so; and the *sense* of ownership could arise out of, or promote, an idea that people have a legal right of ownership over their own data, and this too is an incorrect assumption.

The notion of patient *ownership* of data is misguided, however, both in law and, arguably, in terms of understanding the nature of just what kind of thing medical information is. The DPA gives data subjects the right to *access and inspect* data about themselves and, for example, to request corrections of that data from the relevant authorities where it is inaccurate. However, it accords no ownership rights to patients over information held about them. Indeed, it is perhaps telling that we refer to *patient* data, not *patients'* data. As Lowrance (2002: p3) points out, "there are a variety of medical and legal reasons why no health provider can relinquish possession of, or right of control over, data it has collected in providing or paying for care". While a centre that services cars will inform car owners of information about repairs needed or made to the cars, for legal, financial and quality assurance reasons, the garage must retain and control that data in its own records for a number of years. So, as Lowrance (2002: p4) puts it, "the car owner has a right to be aware of the data but does not *own* the data". So it is for patients: they have a right to be aware of the data but they do not own the data.

The case of ownership seems clear in the case of hospital paper records. Such records are evidently the physical property of the relevant hospital: taking them away without authorisation would be considered theft. Moreover, for reasons of audit, patient safety and administration, it would not be appropriate for a hospital to simply relinquish files of records to patients on demand, given the responsibilities the hospital is under for carrying out functions such as audits and for which such records are essential (and, it might be added, functions that are essential for ensuring provision of the safest and most effective services to *patients*).

The same ownership rights hold for electronic data, although the *sense* of ownership with regard to these is perhaps blurred by the widespread *access* to electronic information of various sorts that individuals now enjoy. Indeed, there are examples of patients whose GPs supply them with CD-ROMs of their medical notes to take home with them: holding in one's hand a physical object could certainly add to the sense that a patient in some way owns not only the object itself but also the information stored on it. Add to this a collective fearfulness among the public in general, and perhaps

among patients in particular, about the security of electronic information, and an anxious desire to exercise control over it becomes easier to understand.

Indeed, it seems that concerns about control over information may often be, at least in part, concerns about data *security*, rather than about privacy or confidentiality. Data security is comprised of measures to maintain controls over access, use and disclosure of information, via physical protections, via contractual restrictions or, in the case of electronic information, via encryption and password protections. Lapses of data security of course do occur, and the media are rarely slow to inform the public about them. People may have reasonable cause for concern on hearing news of misplaced laptops and lost computer disks and perhaps if data security could be completely guaranteed, the desire to have ‘control’ over personal information might be substantially assuaged. But such a guarantee is impossible, for no matter how sophisticated the computer systems or how careful and conscientious the humans who handle our data, there will always be the possibility of technical malfunction or human error. This possibility of error and the risk that it carries in terms of data security is, then, one that individual data subjects must always bear. However, the extent to which it is reasonable to ask them to do so is debatable, and must be weighed against other factors, including the collective benefits generated by the research, which we consider below.

While ownership of patient information may be incorrect in legal fact, nonetheless the existence of a sense among the public of health data being ‘mine’ in some sense cannot be ignored. Responses to a Department of Health consultation on patient perceptions of secondary uses of data demonstrate some evidence of this sense: “It’s my data – information should not be used for secondary uses without my consent, even if it is anonymised”; “The information ‘belongs’ to the patient who has allowed medical staff to have it as part of their care. Once given to medical staff for this purpose, it should not pass out of the patient’s ownership for whatever reason, however good... Allowing others to use patient’s information without their permission is unethical and does not have any part in a ‘patient led’ NHS” (DH, 2010: p27). One reason for this perception of ownership may be linked to certain misconceptions about the nature of medical information and, in particular, the means by which it is *produced*. Health information of any kind, from blood pressure readings to cancer diagnoses to prescriptions, only exists as a result of the input of time and expertise of medical professionals, the use of laboratory testing processes, the use of diagnostic machinery and so on. Indeed, in the UK, most health information only exists as a result of the input of NHS resources of one description or another. So it seems that it may be not only legally but also *conceptually* mistaken to hold to a notion of patient ownership over health information.

By encouraging the conception of patient data as property in some sense, the notion of informational privacy as control also misleads as to what is logically possible in terms of preventing or limiting what others might come to know about us. Information and knowledge are unlike material objects – which we might, logically, possess and fully control – in important ways. First, we cannot control what others come to know about us through accident or inference from information that is freely available to them: as Manson and O’Neill (2007: pp105–6) observe, information can usually be obtained via numerous routes and is inferentially fertile, such that people may draw all sorts of conclusions about us. The point here is that we simply cannot control what people *think* – we might not want to go so far as to suggest that control over information logically implies thought control, but the idea would not be wholly incorrect. Second,

as Manson and O'Neill also suggest, once people know certain information, they cannot will themselves to 'un-know' it. So while I might be able to control whether my friend has a book of mine in her possession – I can demand she give it back or refrain from using it – I cannot demand that she ceases to *know* certain things about me.

The duty of confidentiality

The focus on informational privacy as control over personal information, which is fostered by the DPA, can lead to a view of the patient as an isolated individual making choices about the use of their information in a vacuum. As a result, it can obscure the informational *obligations* that correspond to those informational *rights*, and which take the form of duties of confidentiality. Ethically and legally speaking, confidentiality expresses the relational dimension of privacy: obligations of confidentiality hold in a relationship between two persons, where the confider permits the confidant to gain private – that is, non-public – knowledge about the confider. The confidant in this relationship is under obligation not to disclose the information entrusted to them by the confider and not to use it against them – to harm or disadvantage them.

Where rights of informational privacy focus on the individual's ability to control the use of 'personal' information, informational obligations highlight the specific relational contexts in which information of any sort might be shared and the way in which that information is protected and privacy assured. While a confider gains some control over information in so far as they gain a reasonable expectation that it will not be communicated to others without their consent and that it will not be used to harm them in any way, nonetheless informational obligations do not offer any guarantee of control over information by the data subject. Further, where informational rights offer little purchase on issues of data security, obligations of confidentiality offer an avenue through which to hold data holders to account for the care with which they handle information.

One of the concerns evident from the public debate on the issue of confidentiality in the use of patient information (see, for example, the written and oral evidence given to the Health Select Committee inquiry into the Electronic Patient Record: Health Select Committee, 2007) is that any actual or perceived weakening in the duty of confidentiality could have an impact on trust in the medical profession and on patient behaviour. The advent of the Electronic Patient Record has already raised concerns in this regard, as doctors will usually be the ones who are 'uploading' data to national databases, and so patients who are concerned about how their information will be used on those databases may withhold relevant medical information from their GP, use strategies such as 'doctor-hopping' (changing doctor for different medical events) or even give false names and addresses to prevent a coherent medical record being established. Clearly, a part of this concern is around security of data once it is uploaded, but another part pertains to the impact on the quality of the doctor-patient relationship whereby trust in the confidentiality of that relationship is diminished and 'privacy protective' behaviour adopted because the protection offered by confidentiality is no longer felt to be secure.

If patients are deterred from seeking medical attention, that would likely be detrimental to their own health, but it is also recognised that, as GMC (2009: p18) guidance puts it, "there is a clear public good in having a confidential medical service. The fact that people are encouraged to seek advice and treatment, including for

communicable diseases, benefits society as a whole as well as the individual.” There may be an additional worry in this regard from the point of view of universal access to medical services, or at least in *actual* patterns of access if not in technical universality, that there may be certain groups of patients who are likely to be systematically more wary of how their data is used and who might therefore be systematically more likely to adopt ‘privacy protective’ behaviour, for example, patients whose immigration status is uncertain, patients with a criminal record or who are under investigation by the police, patients who are homeless, or groups of patients with ‘sensitive’ medical issues such as sexually transmitted diseases, mental illness or addictions.

Public benefit

The benefits of using health records for both research and health service management are well known. Many advances in medicine and health provision could not be made without the use of patient data, from medical discoveries in cancer, cardiovascular disease and communicable diseases, to improvements in public health, the ability to identify patient populations at risk, to determine the safety of treatments, to assess the usefulness of diagnostic tests and the appropriateness and adequacy of care, and to evaluate services for clinical and cost-effectiveness.

All of these things benefit both individual patients *and* the public: indeed, it is impossible to separate the two since patients simply are members of the public and the non-patient public are ‘always already’ potential patients – and this is especially straightforward in the case of a national health service, accessible to all and funded by the public in the form of taxes. So, just as there is arguably a public as well as an individual interest in privacy and in the protection of norms of confidentiality in the doctor–patient relationship, so there is a strong interest for individual patients as well as ‘the public’ in the achievement of the benefits of medical and health services research.

The question arises as to whether and when these public, collective benefits are sufficient to justify use of data without explicit consent. As we pointed out in Chapter 1, provisions for public interest defences are available in the legal framework but, due to inadequate definition and lack of clarity about what the public interest consists of, such provisions are less often used than they might be.

From a utilitarian point of view, it might seem obvious that public benefits give a reason for altering the status quo. However, in recent years, there has been a reaction, both in theory and in the public at large, against a crude argument of this sort. In a political culture in which individual rights are recognised, the simple assertion of a benefit is often inadequate to justify action that could put those rights in jeopardy. Thus, in evidence given to the Health Select Committee inquiry into the Electronic Patient Record, a number of witnesses argued that, unless data could be *fully* anonymised, explicit patient consent should be required for data to be used for research purposes. This view was advocated by the Royal College of Psychiatrists and the Royal College of Surgeons (Health Select Committee, 2007; written evidence at EV 103 and EV 106, respectively).

What is needed is both more clarity about what constitutes a public benefit and some form of argument by which the obligations arising from rights and the benefits of public action can be brought into balance. In many instances in society, individuals are willing to accept some limitation on their individual rights or interests, including

their privacy, in order for collective benefits to be secured by government in a way that is impossible for individuals acting alone. Law enforcement, tax collection, provision of public services and national security are partially achieved because of a willingness of individuals to allow collection and use of some personal information. Although *all* citizens might not agree with *all* of these activities, or with the many different purposes they serve, the activities provide substantial collective benefits and often ones that are essential for the functioning of society and for the provision of effective services. Some of these benefits depend on the use of information about individuals. This suggests that one way in which we might approach the problem is in terms of reciprocity and the consideration of what would be a fair contribution to ask of individuals in order to secure the public good.

Fairness and reciprocity

In its evidence to the Health Select Committee inquiry into the Electronic Patient Record, the Academy of Medical Sciences argued that the rights of individuals to restrict access to their data should not be seen to outweigh the public benefits of health research – the Academy commented that “[i]t could be maintained that a patient has the right to say ‘use my data to treat me, but not to improve care for others’”. Or more starkly, “use evidence from other people’s data to treat me, but don’t use my data to help them” (Health Select Committee, 2007: p90, at paragraph 270). The implicit principle being appealed to here is one of fairness, understood as generalised reciprocity. One puts into the general pot as the price for being able to draw out resources.

Despite its intuitive appeal the principle of fairness is not straightforward, either in conception or in application. Thus, questions arise as to whether the use of patient data, and the ‘costs’ to privacy that that may bring, is conditional on an assurance of the public benefit that will arise from research. Further, is it conditional on an assurance of particular *types* of benefit?

Successful medical or health services research always brings benefits of some kind. Even where it does not deliver material advancements in treatments or in services, scientific knowledge and understanding of services are improved. So, in terms of reciprocity conceived as generalised exchange, there are collective benefits, if only in the form of increased scientific understanding. There may be questions as to what costs it is reasonable to ask patients to bear if no material advancements are achieved, but the obvious answer to such questions lies in the fact that the results of research cannot be guaranteed ahead of the research being carried out. Therefore, a project that requires access to data of hundreds of cancer patients may deliver a significant step forward in treatment or it may deliver no step forward at all, but rather just an elimination of the usefulness of that particular line of research. However, the outcome cannot be known before the work is done, and therefore the patient data are still needed.

However, there are instances where the benefits of research may be less straightforward. Some benefits may not be evenly distributed or may even be considered harmful to some groups. Research that establishes the propensity of certain patient populations to disease might be an example of this, as members of those populations can feel stigmatised by widespread knowledge of information about the entire group’s health status. This might include, for example, the propensity of African-Caribbean populations to hypertension and diabetes, or of drug users to HIV, or instances of diseases that are almost exclusive to certain groups, such as Jewish people with Tay-

Sachs disease. The benefits of research about these diseases to these populations seems clear, but it is possible to imagine that the attachment of what may be felt to be ‘labels’ of disease to certain groups may be experienced as stigmatising. Indeed, it may even have harmful economic or social effects if that stigma affects how individuals are treated, for example, in employment or in terms of their eligibility for certain forms of insurance. Interestingly, much of this research can be done on an aggregate level and need not use individual patient-level data, thus showing that costs to privacy may not always be the central ones in the cost-benefit balance – the concerns here are likely to be around issues of social justice instead.

There are also questions about the use of ‘sensitive’ data, usually taken to include information about mental illness and sexually transmitted diseases, and the harms that the use of such information may produce. On the one hand, it is often presumed that patients with this information will not want it to be released for research purposes because of the stigma attached to the relevant conditions: for example, information on mental illness and sexually transmitted diseases is currently not going to be routinely uploaded on to the national ‘spine’ database, which will contain information from individual Electronic Patient Records. However, it seems that one of the benefits of making such data available for research is precisely that such conditions might become less stigmatised and better understood, both socially and medically – especially given that ‘stigmatised’ conditions such as mental illness are often much under-researched.

Different questions arise, however, around the collective benefits gained from the use of data in health services research. The context of such research in the UK is significant. The NHS is a publicly funded service aiming to supply comprehensive, high-quality care without financial barriers to access, and which provides a number of benefits of an individual and collective nature. To secure these benefits, citizens have to cooperate in a scheme of collective provision, for without such cooperation there could not be an institutional arrangement through which any of the benefits are generated.

In this context, use of patient data for health services research could be viewed as a requirement of cooperation to produce collective benefits, which is how the Academy of Medical Sciences appears to regard the matter. Research on public perceptions indicates that people have different attitudes towards the use of data for research within the context of the NHS compared to medical or other types of research using patient data. For example, one respondent to a Department of Health consultation commented on the use of anonymised information that “I am not too concerned for health services or NHS research... or NHS audit. I would be very concerned, in fact totally against health-related data being released to other organizations, for example pharmaceutical industry or other non NHS research organizations” (DH, 2010: p28).

Health services research is an important means by which the NHS is able to plan, provide and manage safe, high-quality, clinically and cost-effective patient care. This could not be done without the use of patient data in clinical audits, evaluation of services, safety checks and so on. If a large number of patients refused to allow their data to be used in this way, the ability of the NHS to plan, provide and manage effective and efficient health care would be diminished, and so too its capacity to provide the collective benefits already mentioned. There seems, then, a fairly strong connection between the use of patient data and the provision of collective benefits by the NHS, and a similarly strong argument to say that if use of patient data was denied, those collective benefits would be, to some degree, undermined.

A case of a somewhat different nature is presented by the use of predictive risk models in health services research. This form of analysis can combine information from NHS hospital admissions, outpatient and Accident & Emergency datasets, to identify patients at risk of re-hospitalisation or admission to a care home. Such patients are 'high cost' and so there are obvious advantages in identifying them early on as this enables preventive or pre-emptive care to be offered, potentially averting higher downstream costs of admission to hospital or residential care. Information about the risk status of individuals is passed back to GPs who can then inform patients and offer relevant preventive interventions. Predictive risk models use pseudonymised data from large datasets. Under the terms of the DPA, therefore, the data are likely to be classed as 'personal' information since the data are 'identifiable', however the data may fall under the Schedule 3 class of information processed for 'medical purposes'. Given the size of the datasets, obtaining consent from patients would be prohibitively burdensome.

There is an important distinction to be made between this form of research and other health services or, indeed, medical research, and which may give rise to different normative implications. While the purpose of other forms of research is *impersonal*, having no direct application or benefit to the original data subject, the aim of predictive risk analysis is precisely to discern information about those subjects and to pass that information back to them, along with the offer of preventive intervention. This would appear to put it in the category of specific rather than generalised reciprocity, since I receive direct benefits from the use of my data. However, this presumes that I experience the news that I am a high-risk patient and the offer of preventive intervention as a 'benefit'. While many patients might be glad of such news and want to take full advantage of preventive monitoring or treatment, others may prefer not to know, particularly given that what is in question is not certain *fact* but rather *risk*, which, by its very nature, is uncertain. Of course, patients routinely undergo tests that may show them to be at 'risk' of some health complication or other. However, the difference between receiving the results of these tests and the results of predictive risk analysis is that a patient would have given their consent to the former and therefore also to being told the results of those tests, but no consent has been given for use of the patient's data in predictive risk research – the patient is therefore in effect receiving results of tests they never knew they had.

Summary

Consent

- Consent acts as a waiver to permit use of data that may otherwise be impermissible.
- But problems arise from the propositional nature of consent and the circumstances in which it is often sought.
- Viewing consent as a form of control over information can ignore these problems.

Privacy

- Informational privacy seeks to protect individuals from various social, psychological and personal harms brought about by unwanted use of information.
- However, delimiting the scope of informational privacy is difficult, and the DPA term 'personal information' is inadequately defined.

- Describing information as ‘personal’ and a subject of privacy protections implies that use of information will affect us *personally* – and in the context of research, that is misleading.

Autonomy

- Autonomy is often understood as implying control over one’s life, and it is therefore linked to the idea of consent as control, as well as to notions of privacy and respect for persons.
- The concept of autonomy underlies the ‘consent or anonymise’ interpretation of the law.

Property rights

- Information is sometimes thought of as a form of property. However, the idea of patient ownership of medical data is misguided both legally and conceptually.
- Questions of data *access* and data *security* blur the edges of the idea of informational privacy and may feed misconceptions about ownership.

The duty of confidentiality

- Confidentiality expresses the relational dimension of privacy.
- It helps ensure relationships of trust between patients and medical professionals, and there are concerns that any perceived weakening of the duty of confidentiality in handling patient information could damage that trust.

Public benefit

- There are clear benefits of research to the public and to individual patients – and to the public as potential patients.
- The question of whether and when public benefits are sufficient to justify use of data without consent raises issues of individual rights.
- Individuals are often willing to accept limitations on such rights in other contexts in order to secure collective benefits.

Fairness and reciprocity

- Fairness as reciprocity suggests that accepting some limitations on individual rights might be a necessary condition for the attainment of public benefits.
- Such acceptance may be conditional on assurance of certain types of benefit, or on particular contexts.
- The NHS provides a context in which citizens have to cooperate in a scheme of collective provision. Accepting the use of one’s data in NHS health services research could be seen as a necessary element of that cooperation.

4. Conclusion

The dominant interpretation of the legal framework that regulates the use of patient information in medical and health services research is, we have suggested, one based around the idea of ‘consent or anonymise’, where consent must be given to use data if it is not fully anonymised. In this approach, informational privacy confers rights of control, and consent acts as a mechanism of that control. This approach seeks to protect individual privacy via the exercise of autonomy over information, and may thereby be viewed as one way of respecting people.

Whatever one thinks of this interpretation, the emphasis it places on explicit consent as a mechanism of control leads to practical difficulties for researchers but also ignores problems around the propositional nature of consent and the circumstances in which it is sought, which make it a highly imperfect instrument of control, even where such control is desired. Additionally, the analogy to control seems too easily made: it overlooks relevant facts about the nature of medical information, elides ideas of access to information with those of data ownership and leads to exaggerated notions of what it is possible to prevent others coming to know about us.

However, contrary to the ‘consent or anonymise’ interpretation of the legal framework, the options of obtaining consent or anonymising data are not the only ones available to legitimate the use of patient information in research. Justifications based on the public interest in research or furthering ‘medical purposes’ are available in law but are inadequately defined and, as a result, researchers may have been reluctant to use them. We suggest that consent as a mechanism of legitimation for use of data ‘kicks in’ too automatically – that is, wherever anonymisation is not feasible – and too quickly – that is, before consideration is given to the possible justifications offered by the public interest in research. Table 1 contrasts the current ‘consent or anonymise’ interpretation with an intermediate ‘controlled access’ approach and, at the other end of the spectrum, a ‘free access’ position.

Conceptually and practically, in our view, the current governance framework over-protects individual privacy – indeed, protects it before it is even threatened – under-privileges the public benefits of research and neglects considerations of fairness and reciprocity, which draw attention to the need to balance individual rights and collective obligations. There is an important question, then, as to whether the current approach to information governance addresses all the social values that are relevant – Table 2 summarises the relevant points. Given this and the problems that prevail around the notion of informational privacy as control and use of consent as a mechanism of control, it seems to us that consideration of the intermediate policy position, as indicated in Table 1, could offer a way to a more soundly reasoned and more reasonable approach to determining the legitimacy of data use. Such an approach might balance the relevant values, avoid automatically imputing privacy-oriented attitudes to patients and the public, and offer more sensitivity to the different contexts and purposes of research.

Box 1: Examples of research using data linkage

Counter-control: consent or anonymise

Key idea is that data subject has veto over use of 'personal', identifiable data. Data cannot be used without consent, apart from under a limited number of conditions where over-ride is allowed. This is a strictly opt-*in* model.

Intermediate position: controlled access

Key idea is that greater access to data is allowed while retaining some consent-based restrictions. Data use may vary depending on context – for example, NHS health services research – or on the distribution or quality of collective benefits. Opt-in or opt-out arrangements may prevail depending on the context.

Free access for ethical research

Key idea is that it is legitimate for bona fide researchers, on ethically approved research projects, to use data without explicit consent, provided the use of the data will not result in harm to the data subject.

Table 2: Summary of social values and their implications

	Counter-control: consent or anonymise	Intermediate position: controlled access	Free access for ethical research
Consent	Consent as a central mechanism is justified by values of privacy and consent, and it should take an explicit and informed structure. Identifiability is key.	Identifiability is insufficient to justify explicit consent, although there will be occasions when consent is necessary.	Explicit consent is not a necessary mechanism, provided that proper professional and research protocols are in place.
Privacy	Strong concern about vulnerability of patient where data is not subject to informed consent requirements.	Concern about patient vulnerability but no <i>requirement</i> of consent, provided other instruments of control are in place.	Assumption that existing codes of practice and professional ethics are sufficient to avoid worries about patient vulnerability.
Autonomy	Autonomy grounds a strong account of explicit and informed consent.	Respect for persons is wider than a concern for autonomy and in any case autonomy does not always imply need for consent.	Research information is not 'personal' in the required sense, and so issues of personal autonomy do not arise in respect of research.
Property rights	Sometimes assumed to imply informed consent, but this assumption cannot be justified.	No assumption that property rights are relevant.	No assumption that property rights are relevant.
Duty of confidentiality	Duty of confidentiality would be breached without consent. Worry about loss of trust in doctor–patient relationship and loss of confidence in system of data collection.	Duty of confidentiality would only be breached under certain circumstances. Worry about loss of trust in doctor–patient relationship and loss of confidence in system of data collection.	No issue of confidentiality arises specific to research. Opt-out system assumes patient agreement to clinicians passing on data. A general concern about loss of trust in doctor–patient relationship in system of data collection remains.
Public benefit	Public benefit will not by itself be sufficient to over-ride obligation to obtain consent.	Public benefit will sometimes be sufficient to over-ride obligation to obtain consent.	Public benefit will always be sufficient to over-ride obligation to obtain consent.
Fairness and reciprocity	Claims of reciprocity weakened because benefit is not sufficiently personal.	Claims of reciprocity sufficient to provide a presumption that people should allow data use in some circumstances.	Claims of reciprocity imply a duty to allow data use.
Summary	Priority of privacy and autonomy, together with concerns about confidentiality, suggest a regime of counter-control through explicit consent.	Concern about privacy, autonomy and confidentiality, but balanced by the public benefit that research creates. Fairness is central in striking the balance – and that balance may shift depending on context.	Public benefit and fairness suggest that autonomy and consent are less important than the public good of research, and privacy thought to be less of an issue, given the impersonal nature of data use.

Bibliography and references

- Academy of Medical Sciences (2006) *Personal Data for Public Good: Using health information in medical research*. London: Academy of Medical Sciences. www.bma.org.uk/images/secondaryusespatientidentifiableinformation_tcm41-169572.pdf.
- Al-Shahi R (2000) 'Using patient-identifiable data for observational research and audit', *British Medical Journal* 321, 1031–2.
- Annas G (2003) 'HIPAA: a new era for medical privacy?', *New England Journal of Medicine* 348, 15.
- Barrett G, Cassell J, Peacock J and Coleman P (2006) 'National survey of British public's views on use of identifiable medical data by the National Cancer Registry', *British Medical Journal* 7549, 1068.
- Benn S (1971) 'Privacy, freedom, and respect for persons' in Pennock, JR and Chapman, J (eds) *Nomos XIII: Privacy*. Silicon Valley, CA: Atherton Press.
- BMA (British Medical Association) (2007) *Guidance on Secondary Uses of Patient Information*. London: BMA.
- 'Campbell v Mirror Group Newspapers' (2004) 2 *All England Law Reports* 995.
- Charity Commission (2011) *Public Benefit: Emerging Findings 2009–2011* (London: Charity Commission) www.charitycommission.gov.uk/Charity_requirements_guidance/Charity_essentials/Public_benefit/assessemmerge2.aspx
- Confidentiality and Security Advisory Group for Scotland (2002) *Protecting Patient Confidentiality*. www.sehd.scot.nhs/publications/ppcr/ppcr.pdf.
- Cox P, Roberts L, Wilson S, Evans B, Ramsay CN, Al-Shahi R and Warlow, C (2001) 'Using patient identifiable data without consent' *British Medical Journal* 322, 858.
- Data Protection Act 1998*. London: OPSI.
- DH (Department of Health) (1997) *Caldicott Committee: Report on the review of patient-identifiable information*. London: DH.
- DH (2003) *Confidentiality: NHS Code of Practice*. London: DH.
- DH (2004) *The NHS Improvement Plan: Putting people at the heart of public services*. London: The Stationery Office.
- DH (2010) *Summary of Responses to the Consultation on the Additional Uses of Patient Data*. London: DH. www.dh.gov.uk/en/Consultations/Responsestoconsultations/DH_109310. Accessed 11 February 2010.
- Dworkin R (1993) *Life's Dominion*. London: HarperCollins.
- European Commission Data Protection Working Party (2007) *Working Document on the Processing of Personal Data Relating to Health in Electronic Health Records (EHR)*. EU Ref. WP131 (Brussels: European Commission). http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp131_en.pdf.
- Fried C (1968) 'Privacy', *Yale Law Journal* 77(3), 475–93.
- GMC (General Medical Council) (2009) *Confidentiality*. London: GMC.
- Godard B, Schmidtke J, Cassiman J and Ayme S (2003) 'Data storage and DNA banking for biomedical research: informed consent, confidentiality, quality issues, ownership, return of benefits: a professional perspective', *European Journal of Human Genetics* 11(suppl 2), S88–S122.
- Haynes CL, Cook GA and Jones MA (2007) 'Legal and ethical considerations in processing patient-identifiable data without patient consent: lessons learnt from developing a disease register', *Journal of Medical Ethics* 33, 302–7.
- Health and Social Care Act 2001*. London: OPSI.
- Health Select Committee (2007) *The Electronic Patient Record, Sixth Report of Session 2006–07*. London: The Stationery Office.
- HMSO (2002) *The Health Service (Control of Patient Information) Regulations 2002*. London: HMSO.
- Hohfeld W (1913) 'Some fundamental legal conceptions as applied in judicial reasoning', *Yale Law Journal* 23(16), 28–59.
- Home Office (2003) *The Legal Framework: The common law*. London: Home Office. www.crimereduction.homeoffice.gov.uk/infosharing22-1.htm. Accessed 4 January 2010.
- Human Genetics Commission (2002) *Inside Information: Balancing interests in the use of personal genetic data*. London: Department of Health.
- Human Rights Act 1998*. London: OPSI.
- Information Commissioner's Office (1998) *Data Protection Act 1998: Legal Guidance*. London: OPSI.
- Information Commissioner's Office (2002) *Use and Disclosure of Health Data: Guidance on the application of the Data Protection Act 1998*. London: OPSI.

- Information Commissioner's Office (2008) *Freedom of Information Act: Awareness guidance 2: information provided in confidence*. Wilmslow: ICO. www.ico.gov.uk/upload/documents/library/freedom_of_information/detailed_specialist_guides/confidentialinformation_v4.pdf. Accessed 5 December 2009.
- Laurie G (2002) *Genetic Privacy*. Cambridge: Cambridge University Press.
- Lewis G, Georghiou T and Bardsley M (2008) 'Developing a model to predict the use of social care', *Journal of Care Services Management* 3(2), 164–75.
- Lord Falconer of Thornton (2001) 'Privacy law and medical research', *The Times*, 17 May.
- Lowrance W (2002) *Learning from Experience: Privacy and the secondary use of data in health research*. London: Nuffield Trust.
- Manson N and O' Neill O (2007) *Rethinking Informed Consent in Bioethics*. Cambridge: Cambridge University Press.
- MRC (Medical Research Council) (2000) *Personal Information in Medical Research*. London: MRC.
- Murray G, Lawrence A and Boyd J (2000) *Linkage of Hospital Episode Statistics (HES) Data to Office for National Statistics (ONS) Mortality Records*. London: OPSI.
- National Health Service Act 2006*. London: OPSI.
- NICE (National Institute for Health and Clinical Excellence) (2008) *Social Value Judgements*. London: NICE. www.nice.org.uk/media/C18/30/SVJ2PUBLICATION2008.pdf.
- Nuffield Council on Bioethics (2010) *Medical Profiling and Online Medicine: The ethics of personalised health care in a consumer age*. London: Nuffield Council on Bioethics. www.nuffieldbioethics.org/personalised-healthcare-0. Accessed 25 October 2010.
- O'Neill O (2003) 'Some limits of informed consent' *Journal of Medical Ethics* 29, 4–7.
- Peto J, Fletcher O and Gilham C (2004) 'Data protection, informed consent and research: medical research suffers because of pointless obstacles' (editorial), *British Medical Journal* 328, 1029–30.
- PRIVIREAL Project *Recommendations from PRIVIREAL to the European Commission*. www.privireal.org/content/recommendations. Accessed 20 June 2010.
- 'R v Department of Health, ex parte Source Informatics Ltd (1999)', *All England Law Reports* (1 All ER), 786–801. London: Court of Appeal, Civil Division.
- Raab C (2002) 'Privacy in the public interest', *The Times*, 21 September.
- Royal College of Physicians Committee on Ethical Issues in Medicine (1994) 'Independent ethical review of studies involving personal medical records', *Journal of the Royal College of Physicians* 2, 439–43.
- Smith G, Shah I, White I, Pell J and Dobbie R (2007) 'Previous preeclampsia, preterm delivery, and delivery of a small for gestational age infant and the risk of unexplained stillbirth in the second pregnancy: a retrospective cohort study, Scotland, 1992–2001', *American Journal of Epidemiology*, 165(2), 194–202.
- Strobl J, Cave E and Walley T (2000) 'Data protection legislation: interpretation and barriers to research', *British Medical Journal* 321, 890–2.
- UKCRC (United Kingdom Clinical Research Collaboration) (2007) *Research and Development Advisory Group to Connecting for Health: Report of research simulations*. London: UKCRC.

Appendix 1: Further details and discussion of legal instruments

The Data Protection Act 1998, principles and schedules

Schedule 1 of the Data Protection Act (DPA) sets out eight data protection principles that apply when processing personal data. Personal data must:

- 1 be fairly and lawfully processed
- 2 be processed for specified, lawful and limited purposes
- 3 be adequate, relevant and not excessive in relation to the purposes
- 4 be kept accurate, and where necessary, kept up to date
- 5 not be kept longer than necessary
- 6 be processed in accordance with data subjects' rights
- 7 be kept secure against unauthorised or unlawful processing
- 8 not be transferred to countries not ensuring adequate data protection.

Schedules 2 and 3 of the DPA set out conditions relevant to the first data protection principle concerning fair and lawful processing. Processing of personal data must comply with at least one condition in Schedule 2, and processing of personal, *sensitive* data must comply with at least one condition in Schedule 3. The relevant parts of those schedules are as follows:

Schedule 2

- 4 The processing is necessary in order to protect the vital interests of the data subject.
- 5 The processing is necessary:
 - (d) for the exercise of any other functions of a public nature exercised in the public interest by any person.
- 6 (1) The processing is necessary for the purposes of legitimate interests pursued by the data controller or by the third party or parties to whom the data are disclosed, except where the processing is unwarranted in any particular case because of prejudice to the rights and freedoms or legitimate interests of the data subject.

Schedule 3

- 3 The processing is necessary:
 - (a) in order to protect the vital interests of the data subject or another person, in a case where:
 - (i) consent cannot be given by or on behalf of the data subject, or

(ii) the data controller cannot reasonably be expected to obtain the consent of the data subject, or

(b) in order to protect the vital interests of another person, in a case where consent by or on behalf of the data subject has been unreasonably withheld.

8 (1) The processing is necessary for medical purposes and is undertaken by:

(a) a health professional (as defined in section 69 of the Act); or

(b) a person who owes a duty of confidentiality which is equivalent to that which would arise if that person were a health professional.

8 (2) In this paragraph ‘medical purposes’ includes the purposes of preventative medicine, medical diagnosis, medical research, the provision of care and treatment and the management of healthcare services.

Section 251 of the National Health Service Act 2006

In England and Wales, Section 251 of the National Health Service Act 2006 (replacing Section 60 of the Health and Social Care Act 2001) provides a power to ensure that patient-identifiable information (‘confidential patient information’ in the wording of the Act) needed for ‘medical purposes’, defined as including medical research and the management of health and social care services, can be used without the consent of patients. There is a sense in which this provision duplicates that available in the DPA, Schedule 3 (8(2)), but provides certainty as to whether or not research meets the relevant criteria.

The Section 251 power can be used only to support medical purposes that are in the interests of patients or the wider public, where obtaining consent is not feasible and where anonymised information will not suffice. Section 251 support is a transitional measure intended to allow time for the introduction of policies to obtain consent and/or techniques for anonymising data. It is reviewed annually, so once researchers decide to apply for a Section 251 exemption support, they must annually justify the use of patient-identifiable data without patient consent to the National Information Governance Board (NIGB), which considers and monitors Section 251 applications. This can leave research activities in a vulnerable position, because arrangements can change dependent on reassessment by the NIGB.

The effect of the original Section 60 (replaced by Section 251 of the National Health Service Act 2006) and the regulations that brought it into force was in effect to modify the common law duty of confidence. The regulations make this particularly clear, and Regulation 4 states that “[a]nything done by a person that is necessary for the purpose of processing confidential information in accordance with these Regulations shall be taken to be lawfully done despite any obligation of confidence owed by that person in respect of it” (HMSO, 2002).

The common law duty of confidentiality

Like all common law duties, the duty of confidence is derived from case law, but it is now echoed in the provisions of the DPA and also established in many codes of professional ethics. A duty of confidence arises “when one person (the ‘confidant’) is provided with information by another (the ‘confider’) in the expectation that

the information will only be used or disclosed in accordance with the wishes of the confider” (Information Commissioner’s Office, 2008: p2).

Compliance with the DPA does not, however, negate the common law duty of confidentiality: when information of a confidential nature is disclosed to someone in confidence, as exemplified by the doctor–patient relationship, then the confidant (that is, the doctor) normally has a duty not to disclose this information without the consent of the confider (that is, the patient). The NHS Code of Practice on Confidentiality summarises the common law as follows: “The key principle of the duty of confidence is that information confided should not be used or disclosed further in an identifiable form, except as originally understood by the confider, or with his or her permission” (DH, 2003: p34). This includes information that the confider would reasonably expect to be considered private.⁵

While there is an obligation on the part of the confidant not to take ‘unfair advantage’ of information imparted in confidence, some government guidance suggests that the purpose for which the information may be used need not necessarily be that for which it was provided, even in circumstances that fall short of those where there is a legal requirement to break confidence or where there is an overriding duty to the public. Home Office (2003) guidance, for example, suggests that “a breach of confidence will only occur where the disclosure of information is an abuse or unconscionable to a reasonable man” (see also *R v Department of Health, ex parte Source Informatics Ltd*, 1999).

Obligations of confidentiality are not, then, absolute. As with exemptions from the DPA, the consent of the confider is only one of the justifications for disclosure of data that would otherwise be considered as a breach of confidence. Other justifications are constituted by legal requirements such as the duty of notification of certain diseases to public health authorities or, as with the DPA, where there is a public interest. The ‘public interest’ is judged according to whether a breach of confidentiality is necessary and proportional, according to the Human Rights Act 1998. In these terms, interference in privacy is seen as justifiable if its object is a ‘legitimate public purpose’ and if the benefits to the public are proportionate to that interference. It has been argued that the protection of health counts as a legitimate public purpose, and that therefore medical research pursuant to that protection can justify breach of confidence (see Academy of Medical Sciences, 2006).

A drawback to the public interest defence is that it is vague and only determined on a case-by-case basis, so its applicability is difficult to predict in advance – it therefore invites cautious interpretation by researchers and health bodies. However, if a research project has been given approval by an appropriately constituted Research Ethics Committee, then it should by definition be in the public interest as well as being proportional and necessary: if research is not in the public interest or involves unnecessary or disproportional interference in privacy, it simply should not have been approved. Section 251 of the National Health Service Act 2006 (previously Section 60 of the Health and Social Care Act 2001) is designed as a mechanism for setting aside

5. The recent case of *Campbell v Mirror Group Newspapers* (2004) tested this notion of information, which the confider could reasonably expect to be considered private and which she would therefore expect to be kept confidential.

obligations of confidentiality on the grounds of public interest, and so can provide added reassurance if litigation for inappropriate breach of confidence is feared. It appears to operate in parallel with the common law duty of confidence, however, rather than to replace it. As such, whether researchers make Section 251 applications is discretionary and dependent on how certain they are of the justifiability of their projects according to the common law of confidentiality.

Human Rights Act 1998

The right to privacy is enshrined in Article 8(1) of the Human Rights Act 1998, which protects the individual's "right to respect for his private and family life". However, this right is qualified by Article 8(2) of the Act, which states that there shall be no interference in the right to privacy except where such interference is "necessary" to protect certain public interests, among which are the protection of health and the protection of the rights and freedoms of others. While Article 8(2) provides for exceptions to the right to privacy, that right is viewed as a fundamental human right and the law views activities to promote and protect human health as being of lesser importance. So, if privacy is thought to be under serious threat, the health benefits of medical research are what must give way, not the other way around. Article 8 rights are qualified rights. This means that in certain circumstances they can be set aside by the state. However, this interference must be lawful, for a legitimate social aim and necessary to achieve that aim.

The issue of proportionality is key here. To merit exceptions to the right to privacy, researchers must establish that the public interest in research meets a "pressing social need" and "is no greater than is proportionate to the legitimate aim pursued" (Human Rights Act 1998, Article 8 (2)). This is judged by a number of factors: the type and amount of personal information used, the number of recipients and the presence of safeguards to maintain security where consent is not obtained. Of course, it can be difficult for researchers to determine in advance whether their research will provide *results* that meet a pressing social need, but it seems likely that at least some information about *potential* benefits will be available.

It is also conceivable, however, that researchers acting in the public interest may also have rights, although qualified ones, to receive and impart information. This is protected by the right to freedom of expression, set out in Article 10 of the Human Rights Act 1998, and applies for instance to journalists revealing information that is thought to be in the public interest. If this were the case, there may be grounds for restrictions on research to be limited on the basis of interfering with the researchers' right to freedom of expression, and it seems at least possible that rules on consent and anonymisation, the onerousness of which many researchers currently bemoan, could be seen as a disproportionate interference with their Article 10 rights.

Appendix 2: Glossary of terms

Aggregate data are data that have been gathered together from several sources and provide information about whole sets of people. They do not include individual, patient-level data, but summarise numerous pieces of such data into a set of statistics. While it is often impossible to discern the identity of individuals from such data, this is not always the case and aggregate data may still be disclosive of individuals' identities (see definition below).

Anonymised data has been stripped of any elements that would allow data users to identify individual patients. For example, anonymised information would not contain the individual's name, address, telephone number, full date of birth or full postcode. NHS numbers or other unique numbers may only be included in anonymised data if no one has access to the 'key' to trace the identity of the patient using this number, or if that key has been destroyed. However, it is arguably impossible to guarantee absolute anonymity: as such, the realistic aim must be to minimise as much as possible the risks of identification. Neither the DPA nor the common law of confidentiality give a categorical definition of data that can be regarded as anonymised. The most that can be said is that, to be considered anonymised, the information should not fall within the definition of personal data in the DPA, nor confidential or private information under the common law (Sections 2.2.1 and 2.2.5, respectively). This means that key-coded data are considered to be personal data unless the key has been destroyed or put beyond the reach of the person holding the coded dataset.

Data controller is the person who determines the purposes for which any personal data are processed, and who allocates responsibilities for data processing.

Data processing What constitutes data 'processing' in the DPA is wide-ranging and includes acquiring, organising, retrieving or using data, or recording, storing, altering or disclosing data. Indeed, legal guidance on the Act notes that "[t]he definition (of processing) in the Act is a compendious definition and it is difficult to envisage any action involving data which does not amount to processing within this definition" (Information Commissioner's Office, 1998: p15).

Data processor is the person who processes data as instructed by the data controller. The data processor must be a separate legal entity to the data controller but will be monitored by them. The data processor may be visible to data subjects. The role of processor does not stem from a general role but rather from a position in relation to concrete activities in a specific context and with regard to specific sets of data or operations. Researchers have successfully used an arrangement in which the data controller (for example, a GP practice or NHS trust) engages the researcher as a 'data processor'. Under this arrangement, researchers are permitted to undertake all data processing procedures that may be performed by the data controller and its employees. In effect, this approach brings the researcher within the same legal entity as the GP or health service and allows them to meet the requirements of the DPA.

Disclosive data are data that do not contain personal identifiers but from which the identity of a person could be deduced by a process of elimination, especially, for example, when a sample size is small, a medical condition rare, a geographical area very specific, or where various values in aggregate data are combined. For example, anonymised data showing unusual patterns of an individual's service use over a particular period of time, and within a particular geographical area, could yield the identity of that individual, especially if the patient population of which they were a part was small. A process known as 'barnardisation' can be used to protect the identity of data subjects in cases where aggregate data are considered to be disclosive. Barnardisation is a statistical tool that disguises actual statistics (and thus ensures that individuals cannot be identified), while still providing an indication of the actual numbers contained in the statistics. However, even with this technique, aggregate data can still be considered disclosive.⁶

Linked data are individual, patient-level data from different sources, which are matched together by means of a common identifier. Such data linkage cannot be done with fully anonymised data.

Patient-identifiable information contains personal details such as name and address or NHS number, which would allow someone looking at the information to identify the individual. For the purposes of the DPA, it is classed as *personal data*, as defined below. Patient-identifiable data are used mostly to deliver care and treatment to patients.

Personal data is "data which relates to a living individual who can be identified – a) from those data, or b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller" (Data Protection Act 1998, Section 1 (1)).

Pseudonymised (linked anonymised) data have had all identifiers (personal information such as name, address and so on) removed but have been allocated a code number that enables the data controller to link the data back to the individual via a 'key' which decodes the data. Once coded data are handed to the recipient (for example, a researcher), the data should be effectively anonymous. However, for the purposes of the DPA, pseudonymised data still fall under the Act's definition of personal data, since it remains possible for the data controller to identify individuals through the key that decodes the data. Pseudonymisation allows information about the same individual to be linked in a way that true anonymisation does not.

Sensitive information is often used as a term to highlight the need for extra care in using information about mental health, sexuality and other areas where revealing confidential information is especially likely to cause embarrassment or discrimination. Note that "sensitive personal data" is defined in the DPA as including all information about "physical or mental health or condition, or sexual life". Processing of this data requires compliance not only with the primary principles of the DPA, but also with the provisions in Schedules 2 and 3 of the Act (DPA, Part I, Sections 2 e and f).

6. A recent decision by the Scottish Information Commissioner found that aggregated data, even when barnardised, could still be classed as personal data in light of the DPA. The case surrounded a request for statistics on instances of childhood leukaemia in Dumfries and Galloway, where the number of cases was so small that the relevant agency (the Common Services Agency) had refused to provide the information on the grounds that individuals could be identified. After a House of Lords review, the Scottish Information Commissioner judged that "aggregated statistics will be personal data if, either on their own or, taken together with the other information..., they enable... individuals to whom the data relates to be identified". Further, it was judged that even if the data were barnardised they could still lead to identification, and that the effect of barnardisation on the actual figures did not have the effect of adequately disguising the data. See Decision 021/2005 Michael Collie and the Common Services Agency for the Scottish Health Service, available at www.itspublicknowledge.info/UploadedFiles/Decision021-2005.pdf.

Appendix 3: International comparisons of data protection

In the EU, the 1995 Data Protection Directive (Directive 95/46/EC) restricts access to certain types of data without consent, but permits member states to make exceptions in the case of health research in the public interest. The incorporation of these exceptions has, however, varied between the different domestic legal systems.

In Holland, personal data used for research is exempted from the Dutch Data Protection Act 1998 if the purpose of the research is not targeted at particular individuals and is not deemed likely to cause distress or damage to a data subject. This means that, under the Act, personal data used in this kind of research can be processed without consent (Godard and others, 2003).


In Sweden, the Personal Data Act (PDA) 1998 establishes similar conditions around identifiability and consent to the UK DPA, but it allows exemptions in deference to the principle of public access to official documents. In addition, the Medical Registers Act 1998 covering transfer and use of personal data by disease registries takes precedence over the provisions of the PDA.


The application of Germany's Federal Data Protection Act 1990 is complex, since it draws more and finer distinctions around the *context* of use than elsewhere: application of the Act depends on the status of the data collecting and storing institution (whether it is public, private, federal or state), and contains different permissions for collecting and using data for different purposes (Godard and others, 2003).

In the United States, the Health Insurance Portability and Accountability Act (HIPAA) 1996 is the primary regulator of patient privacy. HIPAA allows use of medical records for research without the authorisation of the subject if the records are pseudonymised or if a privacy/institutional review board permits a waiver. HIPAA requires that all patients be provided with a 'privacy notice' telling them about access to and use of their records, and informing them that their data can be disclosed for uses related to treatment, payment or other 'health care operations', including research, without any additional authorisation being sought (Annas, 2003).

A Nuffield Trust research summary, *Access to Person-level Data in Health Care: Understanding information governance* by Benedict Rumbold, Geraint Lewis and Martin Bardsley, is also available at www.nuffieldtrust.org.uk/publications. The paper summarises the findings of this report and discusses them in the context of current government policy.

For more information about the Nuffield Trust,
including details of our latest research and analysis,
please visit www.nuffieldtrust.org.uk

 Download further copies of this research summary
from www.nuffieldtrust.org.uk/publications

 Subscribe to our newsletter:
www.nuffieldtrust.org.uk/newsletter

 Follow us on Twitter: [Twitter.com/NuffieldTrust](https://twitter.com/NuffieldTrust)

Nuffield Trust is an authoritative
and independent source of
evidence-based research and
policy analysis for improving
health care in the UK

59 New Cavendish Street
London W1G 7LP
Telephone: 020 7631 8450
Facsimile: 020 7631 8451
Email: info@nuffieldtrust.org.uk

 www.nuffieldtrust.org.uk

Published by the Nuffield Trust.
© Nuffield Trust 2011. Not to be reproduced
without permission.