

9th January 2014

Contact

Alison Hall

alison.hall@phgfoundation.org

The linking and use of biological and health data

Response to Nuffield Council on Bioethics consultation

Introduction

The PHG Foundation is a pioneering independent think-tank with a special focus on genomics and other emerging health technologies that can provide more accurate and effective personalised medicine. Our mission is *making science work for health*; we drive delivery of the benefits of new biomedical and genomic knowledge and applications for health by supporting their rapid and effective adoption within existing clinical and public health services to deliver improved care and outcomes. We also aim to promote a social and regulatory environment that is receptive to innovation, without imposing an undue or inequitable public burden

General comments

1. As implied by the mission statement of the PHG Foundation - 'making science work for health' - our focus is **translating** advances in biomedicine and genomics into better population health. Thus we are broadly supportive of individual biodata being used more creatively and more extensively to further population health. In general we welcome the development of novel resources and novel methodologies to foster these advances (such as those noted in the introductory sections of the Nuffield Council's consultation paper). However we are mindful of the potential risks inherent in such a utilitarian approach, which tends to focus on outcomes (such as improved health) rather than process (such as that comprehensive data sharing may only be achieved by compromising individual autonomy interests).
2. In this response, we focus particularly on genetic and genomic data, since this is an area of special expertise of the PHG Foundation. The Foundation has over 15 years' experience of deliberating these issues. Over this time there have been exponential advances in biomedical technologies such that there is a realistic prospect of whole exome sequencing being offered in clinical health services within the next year or so. We are currently engaged in a project examining the ethical, legal, social and organisational implications of next-generation sequencing for the UK NHS. As part of this project we have held two workshops exploring the emerging ethical, legal

we are broadly supportive of individual biodata being used more creatively and more extensively to further population health

it is inappropriate to insist that all instances of use of personal data are supported by an explicit written consent

and social issues, and the interface between research and clinical care. This project has another 9 months to run, however some tentative findings have emerged, which are consistent with our previous work [1]. This suggests the need for **transparency**: to be explicit about how data might be used by the primary data processor and secondary users; and to state in general terms what the risks and benefits of such data use might be.

3. Where there is **uncertainty** about the safeguards that can be placed on the data or its use, or uncertainty about the risks and benefits to the individual, these should be clearly stated. In general, unless the risks are minimal, some form of **consent** should be sought. We consider that it is inappropriate to insist that all instances of use of personal data are supported by an explicit written consent. This is disproportionate and implies a contractual obligation to keep data secure or to take appropriate measures to reduce its sensitivity (such as by anonymisation) which may not be realistic or feasible. In many cases, a broad or generic consent may be appropriate, or consent may be implied by action (such as where notices are placed in a GP's surgery informing patients that their patient data will be used on an anonymised basis for research unless they object).
4. Whilst we are generally supportive of the right of an individual to opt-out of their data being used in ways that might result in undue risks or burdens, (with the test being a subjective rather than an objective one, as is the precedent for other areas of health law in the UK), the right to object to data use should not be absolute. As intimated in the introduction to the Nuffield Council consultation paper, although individual and public rights are sometimes regarded as being directly in opposition to one another, this is simplistic, and legally incorrect.

The belief that personal biomedical data deserves special protection is, in our view, mistaken

Consultation responses

1. Do biomedical data have special significance?

Possible aspects to consider:

- Is it useful (or even possible) to define biomedical data as a distinct class of data? If it is, what are the practical and ethical implications of different ways of defining this class?
- What factors contribute to the belief that personal biomedical data deserve special protection? Does the sensitivity of biomedical data depend entirely on context or do biomedical data have special attributes that make them intrinsically more sensitive than other kinds of data?
- How are changes in the scope of the data in use providing meaningful insights into individual biological variation and health?
- Do some sub-sets of biomedical data (such as genomic data sets) present particular ethical challenges or offer ethically important benefits?
- To what extent should genomic data sets be regarded as belonging to one individual and to what extent should other interests (e.g. of family members sharing genomic sequences) be recognised? What implications might this have for consent to collection of such data, for feedback concerning the data and for its broader use?

5. Our view is that it is not particularly useful to define biomedical data as a distinct class of data. This is partly because as our understanding of the causation of disease evolves, it is clear that disease occurrence and progression is often more complex than previously thought. Thus for many common diseases (so called common complex diseases such as diabetes, heart disease or cancers) a combination of genetic and environmental factors may combine to cause disease. At each stage of this process there may be factors that expedite or slow down cellular changes. At this functional level, it becomes increasingly meaningless to describe a particular class of data as biodata, since so many factors and variables are potentially involved. Thus dietary and lifestyle data (as evidenced by supermarket purchases) may be necessary but not sufficient for diseases to progress.
6. The belief that personal biomedical data deserves special protection is, in our view, mistaken and arises from the legacy of professional relationships conferring a measure of protection over matters under discussion. Thus the doctor had a confidential relationship with his patient (just as did the vicar or lawyer) which conferred special status for the material under discussion. However health care no longer operates exclusively between doctor and patient and consequently we argue that some biodata deserve special protection but that this depends upon context. Similarly, sometimes

genomic data sets may offer some distinctive challenges, but also offer ethically important benefits which cannot be realised in any other way

genetic/genomic data warrant special protection in view of the sensitivity it confers: some genetic/genomic data are potentially so informative, that it can be used to accurately predict future health or ill health. Thus special protection against its use to unfairly discriminate or stigmatise individuals may be justified. As well as being predictive, genetic/genomic data may also be informative for family members, unlike imaging data. However we suggest that it is mistaken to regard all genetic/genomic data as conferring special significance and therefore worthy of protection (i.e. requiring exceptional responses).

7. Currently, we lack the evidence to interpret common genomic variation in individuals, such as that relating to common complex diseases including cardiovascular disease, diabetes and common cancers. For these conditions, our ability to reach meaningful insights into individual biological variation and health is limited. We need to systematically accumulate evidence about this variation across whole populations (including from those groups for whom evidence is currently sparse including healthy individuals and certain ethnic groups). In order to understand these variations, it may be very important to link genomic variation with outcome data (through health records linkage). However the risks can be minimised through understanding the difference between individuation and identification. Thus as highlighted above, genomic data sets may offer some distinctive challenges, but also offer ethically important benefits which cannot be realised in any other way.
8. If individuals are regarded as having rights over their own genomic data, there is potential for the interests of family members to clash. The rights of individuals might collide (for example where family members disagree whether to access genetic testing or not). Therefore, it is not surprising that some jurisdictions have recognised individuals' special rights to control their genetic data, and that jurisdictions use various mechanisms to manage and sometimes resolve these conflicts [2]. Whilst family members might share potentially informative genomic data (particularly if consanguineous), it is also true that all human beings (by virtue of being human) share large proportions (over 99%) of their genomes, and in turn humans share a considerable part of their genome with other species. So in practice it might be difficult to determine what constitutes a sufficient degree of sharing in order to confer any rights to family members. Also, it is unclear what form these rights would take: a right of notification (of risk or potential test or treatment); a right of access to raw data; a right not to be informed of at-risk status; or even a right of veto (to prevent another family member from accessing a test, such as an immature child in some contexts).
9. The implications for collection of data, for feedback and for broader use of that data are profound. In the clinical setting, genetic testing is usually justified on the basis of clinical utility and 'necessity'. In a research setting, the arguments in favour of accessing data may be less compelling, and some commentators have called for a move towards familial comity or consent [3 4]. Whilst this might sometimes be feasible, it is less clear how feedback can be managed to the extended family (especially within existing budgets).

Our preference is for a measured approach that recognises the likely benefits and the possible harms

2. What are the new privacy issues?

Possible aspects to consider:

- Do new information technologies and 'big data' science raise privacy issues that are new in kind or in scale?
- What are the implications for individual anonymity of linking data across large numbers of databases?
- What is the 'public interest' in biomedical data? What benefits do we want to obtain? In what circumstances might the public interest take precedence over individual and minority group interests?
- What are the actual harms we should seek to avoid in using biomedical data (e.g. discrimination, stigmatisation)? What evidence is there of these harms having occurred?
- In what ways does it matter if people's data are used in ways of which they are unaware but that will never affect them?
- How are applications of computer-based technology (e.g. social networking, image sharing, etc.) affecting concepts of privacy, identity and social relatedness? How are related behavioural norms influenced (e.g. willingness to share and publish data)?
- Would it be helpful to treat biomedical data as 'property'?

10. As previously stated, our focus is on biomedical technologies. In our report *'Next Steps in the Sequence'* which reviewed the impact of next generation sequencing technologies on health care services in the UK, we concluded that the use of such technologies raised issues (including privacy issues) that were novel in scale rather than kind. We did not consider the impact of new information technologies and 'big data' science more generally.
11. When data is linked across large numbers of databases, it becomes more difficult to maintain individual anonymity [5, 6]. In genomic research, participants are often provided with guarantees that their identities will be anonymised if the data is shared. However, with data sharing becoming a default (and indeed a condition of funding for much research), these guarantees seem increasingly thin. The difficulties inherent in maintaining individual anonymity are acknowledged by the Personal Genome Project, which requires that its participants place their identified data in the public domain without any attempt at anonymisation or de-identification. Our preference is for a measured approach that recognises the likely benefits and the possible harms, including potential threats to individual anonymity of data linkage, future uncertainty about how the field will develop, and the steps that can currently be taken to mitigate these challenges.
12. Public interest in biomedical data can be interpreted in a variety of ways. One instance of public interest is that described in Article 8(2) of the Human Rights Act 1998 which describes the right to interference with the right to

Discrimination and stigmatisation are frequently cited harms, yet there is scanty (and sometimes anecdotal) evidence of this occurring

respect of an individual's private and family life, home and correspondence – in the following terms:

“There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others”.

13. In the context of health care law, the most commonly cited examples where the public interest has been invoked are in areas of public health (such as infection control); mental health (enforced hospitalisation) and breaching confidentiality where disclosure is justified on the basis that it is likely to avoid serious, imminent harm to the individual making the disclosure or to another person. In genomic research, the claims to public interest are typically less intense. One example is building evidence bases of genetic and genomic variation as a prerequisite to understanding disease causation and progression. Undoubtedly such work is in the public interest in that such knowledge is the foundation of improved health care. However in this context, the case for necessity might be less compelling than in some of the other examples described above.
14. Discrimination and stigmatisation are frequently cited harms, yet there is scanty (and sometimes anecdotal) evidence of this occurring. In the context of genetic testing, fear of genetic test results being recorded in medical records is sometimes cited as a driver for consumers to seek genetic tests on a direct-to-consumer basis rather than through a health care provider. The potential deterrent effect of genetic discrimination on test uptake continues to be cited in support of renewal of the UK Concordat and Moratorium on Genetics and Insurance, which concerns predictive genetic tests.
15. The use of people's data without their knowledge could undermine public trust and cause people to be more cautious about how their data is used. Public trust is hard to establish but easily squandered as evidenced by episodes such as the Alder Hey Inquiry which ultimately resulted in the Human Tissue Act 2004 being implemented. Extending property rights to body parts and to data emanating from the body was something that was debated in the early 2000's particularly following some court cases that seemed to establish such a right. However, in our view, establishing property rights over data does not address some of the most problematic elements described above – where there are multiple claims over the same data (whether these are multiple individual claims from family members or a clash of private and public interests). Another way of addressing these issues might be to consider the intent of the data processor, so that wilful de-identification which is not justified by public interest could be an offence[7].

the proposed draft EU Data Protection Regulation seems likely to impose significant burdens

3. What is the impact of developments in data science and information technology?

Possible aspects to consider:

- To what extent and in what ways has the availability of biomedical data and new techniques for analysing them affected the way in which biomedical research is designed and funded? Is there any evidence that these factors have affected (or are likely to affect) research priorities?
- What are the main interests and incentives driving advances in data science and technology that can be applied to biomedical data? What are the main barriers to development and innovation?
- Does 'big data' need a more precise definition or is it a useful concept in the life sciences even if loosely defined? Has enthusiasm for 'big data' led to over-inflated expectations on the part of governments, researchers and/or the general public?
- What are the significant developments in the linking or use of biomedical data, including any we have not mentioned, to which we should pay attention in our deliberations?

16. Biomedical research is increasingly a global endeavour with multiple collaborators. The requirement for powerful computing capacity supported by a skilled workforce potentially marginalises research done in low and middle income countries with the data flows tending to be from those countries to high income countries. Increasingly, placing research data into a publicly accessible database once the research is completed is a condition of receiving research funding in some jurisdictions.
17. Advances in science and technology often arise from academic research programmes incentivised by the desire to make new discoveries about individual and population biology. However research in the health/life sciences sector has been facilitated by novel technologies emerging from basic computer science, physics and statistics/mathematics research, particularly in areas such as the development of big data analytics for genomic data. This research has then been exploited by IT and web companies and then re-purposed for health purposes. Through this process, the pharmaceutical industry is now able to exploit the tools and knowledge that have arisen from these basic science endeavours and thus commercial interests are increasingly being seen as driving this innovation. Drug development is a slow process, with research and development and ethical review being cited as potential barriers. Other barriers include restrictive data protection regimes, which have the potential to impose excessive or disproportionate safeguards on data sharing and data flow between countries. In this respect, the proposed draft EU Data Protection Regulation seems likely to impose significant burdens on this sector. More generally, a significant barrier to innovation is the lack of resources put into the process

there needs to be investment in the infrastructure, and particularly in the modernisation of IT systems to enable the collection, analysis and use of biomedical data

of translation, particularly within the NHS. In order to implement these changes, there needs to be investment in the infrastructure, and particularly in the modernisation of IT systems to enable the collection, analysis and use of biomedical data in the provision of individual healthcare through the wider adoption of electronic patient records. As an organisation interested in the translational process, this is something that we are particularly keen to highlight.

18. The concept of 'big data' is useful, but there is a tendency for it to be used indiscriminately. It has been used by different stakeholders to serve different agendas, promoting a range of different expectations – some inflated - but using the term more precisely will not necessarily address this.
19. The routine longitudinal linkage of electronic health records with a diverse set of other databases (including lifestyle and dietary information) has the potential to create an evidence base which can be used to more accurately predict how genes and environment combine to impact upon health. Rationalising the sharing of patient records between health care sectors and between health and social care settings will also help to establish more seamless and integrated care, and is therefore an important priority. Establishing robust regulation and governance of such data sharing and creating clear and comprehensive guidance is a prerequisite. The Health and Social Care Information Centre has statutory responsibility under the Health and Social Care Act 2012 to produce a code of practice for processing confidential information from health and social care services in England, and will play a part in implementing the recommendations from the Information Governance Review led by Dame Fiona Caldicott. In addition, the new Clinical Practice Research Database (a joint venture from the MHRA and the NIHR incorporating the existing General Practice Research Datalink and other NIHR research capability pilots, <http://cprd.com/about/>) aims to maximise the health gain that can be achieved through the use of anonymised linked NHS data in research studies and to help improve the way clinical trials of innovative medicines can be undertaken. An additional objective is to gain funding for research projects that increase the wealth of the UK as a whole. The Farr Institute also offers health informatics research expertise (<http://www.farrinstitute.org/>).

Increasingly collaborations call for creative ways of sharing and processing data across jurisdictions with different privacy regimes

4. What are the opportunities for, and the impacts of, the use of linked biomedical data in research?

Possible aspects to consider:

- What are the hopes and expectations associated with data use for biomedical, public health and life sciences research? What are the main concerns or fears?
- To what extent do the kinds of collaborations required for data-driven research (e.g. international or multi-centre collaborations) generate new ethical and social issues and questions to those in other forms of research?
- Should researchers be required to allow others to access data they have collected for further research?
- What sorts of concerns are raised when research is carried out by a commercial firm?

20. The hopes are that the combination of data will yield unprecedented insights into disease causation and therefore catalyse the development of novel treatments, preventive interventions and novel predictive and therapeutic tools; that it will become possible to more accurately predict the risk of ill health in individuals, and ultimately reduce morbidity and mortality in that group; and more generally that this will generate a greater understanding of why some treatments work for only a subset of the population. The fears are that in combination, these data will undermine personal privacy, and promote a world in which individuals are fearful that predictive information about their health might lead to discrimination such that they become unemployable and uninsurable. There are also fears that commercial interests will dominate drug development and drug access more than they do so at present, to the extent that they create an underclass of the vulnerable that will become ever more disenfranchised.
21. Increasingly collaborations call for creative ways of sharing and processing data across jurisdictions with different privacy regimes. This might include cloud computing, which might be less secure than data held in other ways. The legal basis of cloud computing is not yet well worked out. It also remains unclear how responsibilities and duties should be shared between collaborators. For example, if data is processed for the purpose of clinical care, and then passed to a biobank for ongoing storage, and subsequently accessed by secondary researchers who identify an incidental finding of serious treatable disease, it is unclear who has what responsibility to further investigate those findings, or feedback results to the patient, particularly if the only person who has a continuing relationship with the patient is the clinician [8].
22. Researchers should not necessarily allow others to access the data they have collected. This depends upon the terms of the consent gained from the research participant. However in general, we think it is advisable for

Those accessing the research should comply with proportionate safeguards

broad consent for further research to be sought from research participants, on the basis that any research should be ethically approved; that the broad purpose of research is known to the data subject, after explaining any potential risks or burdens. Those accessing the research should comply with proportionate safeguards – such as to keep the data secure, that employees should be subject to contractual requirements to keep data confidential *etc* – enforceable through contractually binding agreements.

23. When commercial companies are involved, concerns may be raised about the potential exploitation of research participants. More generic concerns relate to concerns that data generated may be withheld from publication (particularly if they constitute negative data). There might be concerns that patient data may be misused or shared inappropriately. Additional concerns are that access to drugs or interventions will ultimately be limited, if commercial companies charge excessive amounts, and that exclusivity might drive up prices. Finally, in the context of direct-to-consumer genetic test provision, there might be concerns about test quality, validation and performance.

increased use of linked electronic data might facilitate seamless and integrated care

5. What are the opportunities for, and the impacts of, data linking in medical practice?

Possible aspects to consider:

- What are the main hopes and expectations for medical practice associated with increased use of linked electronic data? What are the main concerns or fears?
- What can be said about public expectations about the use of health care data, in terms of appropriate use, information and control? To what extent would members of the public expect health care data to be shared with other agencies or bodies?
- Is there potential for privacy controls to hide secrets, such as abuse, or to disadvantage people in unintended ways (by preventing best treatment, perhaps)?
- Are there particular issues raised by 'risk-profiling' where individuals at high-risk (e.g. of type 2 diabetes) are identified and approached for specific interventions? What might make the difference between this being intrusive and it being supportive?
- What are the implications of episodes of treatment across different care providers being used routinely as research data? How might this affect the ethical basis of the doctor-patient relationship?
- To what extent does the possibility that biomedical data can contribute to a research base to advance the effective treatment of others create a moral obligation to allow them to be used in this way? What might limit this obligation? How should we regard (and provide for) those who refuse to allow their data to be used?

24. It is hoped that increased use of linked electronic data might facilitate seamless and integrated care between primary, secondary and tertiary sectors. It is hoped that it will promote increased efficiency and effectiveness, reducing the extent of duplication of recording patient details, and improving patient safety (through more systematic drug dosing, and recording of allergies, drug interactions *etc*). Electronic records also provide improved systems for managing tiered access, and an explicit audit trail. The main concerns are that, in practice, electronic systems will not work as effectively as paper records (either because of lack of interoperability of electronic systems between providers, or lack of a particular functionality, such as in clinical genetics, where different subfolders are kept for each family member which allow the provenance of genetic information within the family to be made explicit); that they are inappropriate for some types of record keeping (such as complex familial pedigrees in clinical genetics) and that access based controls will be insufficient in preventing sensitive and sometimes predictive patient data from being accessed inappropriately, with the potential for causing patient distress.

We favour a system in which there is voluntary contribution to a research data base

25. Our organisation does not handle patient data. However, others report that patients expect their data to be shared to facilitate their care, however they may underestimate the extent to which their data is routinely shared, sometimes for purposes that do not appear to be directly relevant to their care (such as audit or financial management).
26. As part of the Connecting for Health initiative, a consultation explored the potential for summary care records to have privacy controls imposed on them through 'sealed envelopes' both by the patient and health care provider. This was not implemented, but provides a precedent for this type of thinking. The sealed envelope system was proposed to address contexts where patient data is judged to be particularly sensitive, such as clinical genetics and infectious diseases, however, this is precisely the context where an individualised approach seems to founder: other examples of where multiple people might have competing claims over data include where two estranged parents disagree over the healthcare needs of a child. The availability of privacy controls does not resolve the underlying ethical tensions imposed by an individualistic approach.
27. The PHG Foundation has explored the ethical, legal and social issues raised by risk-profiling in the context of screening for common cancers as part of the COGS project [9]. We concluded that the most challenging issues were likely to be generated from the manner in which individual genotype data was accessed (i.e. blood samples taken) analysed and stored, rather than participants regarding the process as intrusive or supportive. The aim of the COGS project was to target screening interventions more effectively through the use of genotype data, rather than, as in type 2 diabetes risk-profiling to proactively target currently asymptomatic 'at-risk' individuals. In this case, there are multiple questions raised about the boundaries between health and 'ill-health' and the role of medical professionals in pro-actively identifying those who feel well but who are risk. In this context, key aspects are public education and communicating the health risks and benefits associated with early detection and treatment.
28. Our view is that there should be more systematic recording and analysis of routine health data (including episodes of care) for research. Creating a robust, complete database which can be interrogated for multiple purposes (such as the Clinical Practice Research Datalink CPRD) is a pre-requisite for planning an effective, responsive, flexible and accountable health service. As previously stated, we believe that openness and transparency is the key to securing patient trust in the health services.
29. We favour a system in which there is voluntary contribution to a research data base. Although some have framed this in terms of a reciprocal relationship such that only those who contribute their data to research have a right to access routine healthcare, we feel that this is unduly coercive. It is important to treat people who have reservations with respect, and to provide as full and complete an explanation as possible as to how their data will be used. If they continue to withhold their consent, this view should be respected. They should be warned of any foreseeable impacts upon their care as a result of this decision. Sharing patient data for research should be distinguished from sharing data for service development (such as clinical audit and financial management).

it is difficult to create robust categories of legitimate 'end-users' without creating systems that are bureaucratic, onerous, and costly

6. What are the opportunities for, and the impacts of, using biomedical data outside biomedical research and health care?

Possible aspects to consider:

- What are the main hopes and expectations associated with the wider use of biomedical data (outside biomedical research and clinical practice)? What are the main concerns or fears?
- What factors are relevant to determining the legitimate scope of further uses of biomedical data? For example, should it be restricted to a 'compatible purpose' (and, if so, how might this be defined)? To uses that are in the 'public interest'? To use only by public authorities (and those providing public services under contract)? To non-commercial or non-profit uses/users?
- What are the ethical implications of using predictive analytic tools with biomedical data outside health care and research (e.g. in recruitment or workforce management)?
- Would the ability of individuals to maintain direct control over the use of data about them be likely to affect the range of further uses to which they would allow the data to be put?
- Should individuals be able to profit from the use of their biomedical data (e.g. by selling access to the data to commercial companies)?

30. Our focus is largely the use of genetic and genomic data for health care purposes and for related research. We are also interested in how health data may be accessed by other secondary users including insurers and pension providers to make inferences about individual morbidity and mortality. With these uses in mind we endorse the concept of 'no surprises' to the extent that patients and research participants should understand the potential uses to which their data might be put.
31. Our view is that it is difficult to create robust categories of legitimate 'end users' without creating systems that are bureaucratic, onerous, and costly. Justifying data processing and use, whether in terms of public interest or that the user is non-commercial, seem unduly simplistic and subject to exceptions. We therefore favour a data sharing environment that is more open, but where reasonable efforts are made to ensure that data subjects are familiar with the risks and benefits associated with the wider use of biomedical data for non-medical purposes.
32. In considering broader uses of biomedical data (outside biomedical research and clinical practice), we believe that it is important for governments and regulators to propose legislation that is proportionate and enforceable. In the short-term, creating circumscribed spheres of control of biomedical data seems unenforceable. Since market forces seem likely to dictate the wider use of these data, it seems to us that the ultimate test will be whether biomedical data serves as a good enough predictor of life events (such as

in legal terms,
in the UK
there is a
longstanding
prohibition
against selling
body parts for
gain, on the
basis that it
is contrary to
public policy

chronic ill health, susceptibility to occupational illness, or premature death) to give it more utility than existing data that is routinely collected (such as family history information).

33. Individuals already derive benefits from authorising the use of their data by others (for example through supermarket loyalty card schemes). In some senses, the sale of individual biomedical data does not pose any additional ethical questions, except where that data necessarily reveals information about another individual. However in legal terms, in the UK there is a longstanding prohibition against selling body parts for gain, on the basis that it is contrary to public policy. If individuals were allowed to profit from selling their biomedical data, the distinction between these two approaches should be made more explicit: more empirical work needs to be done to assess public attitudes as to how payment might change the nature of the relationship between patients and health care professionals. Since in some respects, the UK 100,000 Genomes Project includes the planned commercial exploitation of NHS patient records by a limited company wholly owned by the Department of Health (GEL) it also might need to be made explicit why such sales are acceptable for institutions but not for individuals.

We do not support the proposal that linked biomedical data requires distinctive governance arrangements

7. What legal and governance mechanisms might support the ethical linking of biomedical data?

Possible aspects to consider:

- What ethical principles should inform the governance of biomedical data? For example, should the principle of 'respect for persons' be given primacy here? How might this relate to principles such as solidarity and tolerance?
- Does the use of linked biomedical data require distinctive governance arrangements compared to the use of other personal data?
- Are the current principles of consent – including the principle that consent can be withdrawn – still 'fit for purpose' in relation to the linking of biomedical data?
- What level of continuing involvement is it reasonable to expect individuals to have in how their data are used after they have been collected?
- Should there be an opt-in or an opt-out system for people to decide whether to allow their personal medical data to be used for public benefit?
- Under what conditions ought individuals to be content to delegate authorisation of the use of health and biological data about them?
- What role should public engagement and democratic processes play in the determination of governance measures? In what circumstances, if any, might the outcome of democratic procedures mandate overriding individual interests?
- What inconsistencies exist in current ethical guidance and governance structures relating to biomedical data?
- What examples are there of innovative initiatives that promote privacy while encouraging participation?

34. We are supportive of 'respect for persons' as an ethical principle underpinning legal and governance mechanisms in this area, but consider that this principle is not absolute. We do not support the proposal that linked biomedical data requires distinctive governance arrangements on the basis that linkage in itself does not justify special safeguards. Similarly we argue that whilst many principles of consent remain applicable, it is difficult to see how the right to withdraw (or indeed to be forgotten as proposed in the draft EU Data Protection Regulation) is feasible in some instances of data pooling and linkage. However, we are not clear whether it is 'consent' that is not fit for purpose, or that the process of consent is simply being required to do too much. In our opinion a valid consent for the use of linked biomedical data might include making it explicit that a right of withdrawal is only partially actionable (either in time, such as having an initial cancellation period, or in nature - applicable to certain types of data or user).

special exemptions to the requirement for patient consent are given to cancer registries, but not other disease registries which could arguably provide equivalent public benefit

35. The extent of continuing involvement of individuals in how their data is used depends substantially upon context. As previously stated, we feel a provision that limits data usage to activities in the public interest may exclude legitimate uses. We also suggest that providing that use may proceed in the public interest is not synonymous with use of data for public benefit (as stated in this section). Subject to that proviso, we are generally supportive of establishing an opt-in or opt-out system by which individuals can authorise their personal medical data for a range of uses for public benefit, such as ethically approved biomedical research, but without further clarification we regard the default of 'for public benefit' as being too broad a category (if after the Human Rights Act provision it includes use for economic benefit or to prevent crime).
36. The law currently provides delegated authority for those who lack capacity to consent (through age, disability or illness or those with fluctuating capacity) on the grounds of necessity. What is being contemplated in the consultation paper is similar to the grounds on which the exemption for historical and statistical research operates under the Data Protection Act (DPA), namely that the data are not processed to support measures or decisions with respect to particular individuals (DPA 1998 Section 33(1)(a)) and that data are not processed in a way that causes, or is likely to cause substantial damage or distress to any data subject (DPA Section 33(1)(b)). Additionally data processing should be for public benefit. We support a regime that establishes a similar basis for processing linked biomedical data under delegated authority.
37. The principle of deliberative democracy mandates delegation of certain individual decisions to an elected government through the democratic process via published manifestos and the electoral process. During Parliamentary terms, Government consultations provide a mechanism for public engagement. There are many examples where individual interests are overridden on the basis of democratic mandate (such as taxation or road safety): there are many circumstances in biomedicine where democratic processes mandate overriding individual interests (such as surveillance for infection control; enforced hospitalisation *etc*). Legislative bodies, professional organisations and other stakeholders in positions of authority, have a continuing responsibility to engage with publics who may be affected or influenced by decisions that they make.
38. There are many inconsistencies in the governance of biodata particularly between jurisdictions. One of the most glaring inconsistencies concerns the claims that are made about the feasibility of anonymising genomic data. Another is that special exemptions to the requirement for patient consent are given to cancer registries, but not other disease registries which could arguably provide equivalent public benefit. One challenge to good governance is the lack of standardised datasets to inform the interpretation of human and pathogen genomic variants. This issue is likely to become more pressing in the future as novel genomic technologies such as whole genome sequencing (WGS) and whole exome sequencing (WES) become used more widely in public health. This highlights the need to standardise the genomic information that is analysed, reported and retained.

Novel technologies known as participant-centric initiatives employ social media techniques to provide more dynamic and flexible consents

39. A number of innovative initiatives have been developed to promote privacy, and encourage participation from patients and participants. Novel technologies known as participant-centric initiatives employ social media techniques to provide more dynamic and flexible consents [10]. Organisations like patientslikeme.com explicitly use social media to facilitate research participation but other organisations increasingly use social media to increase their visibility.

Bibliography

1. Wright C, Burton H, Hall A, *et al.* Next Steps in the Sequence: The implications of whole genome sequencing for health in the UK. PHG Foundation, 2011.
2. UK Parliament. Human Tissue Act 2004.
3. Davey A, Newson A, O'Leary P. Communication of Genetic Information within Families: The Case for Familial Comity. *Bioethical Inquiry* 2006;3:161-66, doi: 10.1007/s11673-006-9022-5
4. Parker M, Lucassen A. Genetic information: a joint account? *BMJ* 2004;329:165-7.
5. Homer N, Szelinger S, Redman M, *et al.* Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density SNP Genotyping Microarrays. *PLoS Genetics* 2008;4(8), doi: 10.1371/journal.pgen.1000167
6. Gymrek M, McGuire A, Golan D, *et al.* Identifying Personal Genomes by Surname Inference. *Science* 2013;339:321-24, doi: 10.1126/science.1229566.
7. Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data, 2007.
8. Wolf S, Crock B, Van Ness B, *et al.* Managing Incidental findings and research results in genomic research involving biobanks and archived data sets. *Genetics in Medicine* 2012;14(4):361-84, doi: 10.1038/gim.2012.23
9. Hall A, Chowdhury S, Hallowell N, *et al.* Implementing risk-stratified screening for common cancers: a review of potential ethical, legal and social issues. *Journal of Public Health* 2013:1-7, doi: 10.1093/pubmed/fdt078
10. Kaye J, Curren L, Anderson N, *et al.* From patients to partners: participant-centric initiatives in biomedical research. *Nature Reviews Genetics* 2012;13:371-76, doi: 10.1038/nrg3218].