

Call for Evidence on the Current Data Protection Legislative Framework Response from the PHG Foundation

Introduction

The Foundation for Genomics and Population Health (PHG Foundation) is the successor body to the UK Public Health Genetics Unit. Its overarching purpose is to foster and enable the application of biomedical science, particularly genome-based technologies, for the benefit of human health. Among its specific objectives is the promotion of a social and regulatory environment that is receptive to innovation, without imposing an undue or inequitable public burden. The Foundation has a particular interest in the way that new technologies are translated within health services, in genetic research and its impact upon clinical and public health services.

Relevant experience in this area

The Foundation's remit is as a policy think-tank with the overarching aim of making science work for health. The PHG Foundation has experience in the area of data protection, having contributed to the Academy of Medical Sciences report 'Personal data for public good: using health information in medical research' (2006) (AMS report).¹ In this consultation response we focus upon a small number of consultation questions which have particular relevance to our work, and from which general principles can be drawn rather than seeking to provide evidence of data processing practice (as we do not process data from large numbers of research participants or individuals).

In the following sections we make some general points concerning:

- The nature of genetic data
- The inadequacy of specific consent as a prerequisite for epidemiological research, and
- The nature and linkage of medical records between family members in the context of clinical genetics.

Consultation questions

1. **What are your views on the Data Protection Act 1998 and the European Directive upon which it is based? Do you think they provide sufficient protection in the processing of personal data? Do you have evidence to support your views?**
 - 1.1 There are difficulties in which the Data Protection Act 1998 (DPA) and the Directive on which it was based address the use of personal data for population-based research. Some of these difficulties were noted in the above-mentioned report by the Academy of Medical Sciences. One of the

¹ Academy of Medical Sciences (2006) Personal data for public good: using health information in medical research.

conclusions of this report was that the legislative framework was satisfactory but that the current regulatory framework (rather than the legislation) is damaging research using personal data (at page 21). Research tools used to assess the burden of disease within populations often require the secondary use of personal data and there is a lack of clarity about how this data should be legitimately accessed and processed. Sometimes this data can be anonymised so that data can be processed without securing the consent of individuals. Often however, some element of identifiability is required for useful research to be done. One possible alternative approach that of securing consent may be impractical, prohibitively expensive or compromise the quality of the data (by introducing bias).²

1.2 In 2006 we shared the view that the legislative framework was sound. Since then, the calls for reform of the Data Protection legislation have become increasingly persistent. For example, the Data Sharing Review 2008 by Thomas and Walport called for change claiming that the law and its framework 'lack clarity, responsiveness and bite' arguing that there was a need to 'clarify and simplify the legal framework governing data sharing'.³ We now agree that some modification of the Directive and DPA is necessary, to take account the advances that have taken place in the ways that data is now shared, accessed and processed (including technological advances and social changes such as social networking via the internet, cloud computing etc). We also believe that calls for reform of the DPA are more pressing given that the changes envisaged in the AMS report (such as increased use of the Section 33 DPA exemption) have not come about.

2. What are your views on the definition of 'personal data' as set out in the Directive and the DPA?

Our view is that the current definition of 'personal data' as set out in the Directive and the DPA is quite difficult to apply in the context of genomic data, particularly given the pace of scientific developments in this area. This is for a number of reasons.

2.1 Effective anonymisation of genomic data

Genomic research over the last decade has been characterised by a number of different research methods. One tool (the Genome Wide Association Study) involves the analysis of genomes from large cohorts of individuals typically with and without a disease, to identify genetic variants associated with the condition. Using this tool, many hundreds of genetic susceptibility variants have been identified. These methodologies typically involve collaboration between many research groups and a number of consortia have embraced a policy of providing open access of their research results, to promote research effort in this area. The field was thrown into disarray by the publication of a paper⁴ which described a method by which individuals within a pooled group of samples could be re-identified, thus raising questions for the participants about the adequacy of the consent

² Section 251 Health Services Act support may also be applicable.

³ R Thomas, M. Walport (2008) Data Sharing Review.

⁴ Homer N et al (2008) Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. PLoS Genet 4:e1000167

that had been obtained, the security of the data set and the extent to which this type of work should properly fall under the Data Protection legislation.

2.2 Linkage of different types of data sets

A second feature of advances in genetics and genomics is that in order to combat disease, there are efforts to combine information about genotype, as well as phenotype (including potentially sensitive information about lifestyle such as alcohol intake or socio-economic class). This linkage between different data sets is likely to become a crucial element in both medical research and delivery of health care over the next decade.

2.3 International collaboration

As mentioned above, much genomics research involves international collaboration. Such genetic epidemiology might involve both data and samples being sent across jurisdictions. Whilst greater harmonisation of data processing legislation at European level might be desirable, this might be of limited utility in the context of genomic research effort, since collaborations increasingly include samples from the USA and Asia which have different regulatory frameworks. In the USA for example, genomic research may not be regarded as human research because it falls outside the Health Insurance Portability and Accountability Act (HIPAA).

2.4 Genetic exceptionalism and the extent of 'sensitivity' of genetic data

The emergence of new genomic techniques such as whole genome sequencing, now allows individual genomes to be described in incredible detail. This results in a voluminous list of bases (or components) which make up the genome, including coding and non-coding regions. Whilst this is a remarkable technical feat, it is becoming clear that particularly where such vast amounts of data are generated, that the majority of these data may be mundane and uninformative. Thus the view that genetic or genomic data should be ascribed 'special' status in revised legislation is both incoherent, inconsistent and is not proportionate because only a small amount of genomic data could be regarded as 'sensitive' (given our current state of knowledge).

2.5 Linkage of data relating to family members within a pedigree

There is little recognition within the DPA of the need to link medical health records of family members for clinical reasons. For example in clinical genetics, the usual practice is for health professionals to take a family history and construct a pedigree which will contain personal medical data relating to many family members (including relevant diagnoses, age, and sometimes personal contact details etc). Another complicating factor is that multiple individuals from the same family might be seen individually by the same genetic centre (or referred to another specialist centre if they live outside that geographical area) and separate notes might be held which relate to each individual and to the family as a whole.

7. Are there any other types of personal data that should be included? If so, please provide reasons why they should be classed as 'sensitive personal data'?

There are undoubtedly sensitivities associated with some types of genetic data. For example, genetic data about one individual can sometimes be used to generate clinically useful information about family members; some genetic data can be used in a predictive manner to identify future ill health or susceptibility to future disease; genetic variant data may indicate the efficacy of a drug treatment or the course of a disease. As a result, there are fears that genetic data might form the basis for genetic discrimination or stigmatisation (and examples abound over the last century, for example the treatment of those with sickle cell disease in the USA). For this reason, some have called for genetic data to be given special treatment in revised legislation. Our view, as stated above is that this genetic exceptionalism is unjustified.

33. Should the definition of consent be limited to that in the EU Data Protection Directive i.e. freely given, specific and informed?

33.1 In the context of genomic research and epidemiology, it is possible that the requirement for specific consent from individuals participating in the research might compromise population health. A good example is where a new infectious disease has been described and epidemiology is needed to assess its virulence, spread and identify at-risk groups: (examples include the Ebola virus or swine flu). Increasingly this might involve investigating genetic variants which are associated with the disease. For example, CJD seems to have occurred preferentially within individuals with a particular genotype. If a prerequisite for carrying out this type of research is that a specific consent has been obtained, then this would prevent such research from occurring and would be detrimental to population health.

33.2 In other research contexts, particularly those involving whole genome analysis certain research consortia require some degree of public access to the research data. As this technology is in its infancy, little is known about the potential risks and harms that might be associated with publicising personal information in this way, and obtaining a specific consent to the harms which might arise may be almost impossible (since the harms of disclosure are not yet fully known).

36. Do you have evidence to suggest that the exemptions are fair and working adequately?

36.1 Anecdotal evidence suggests there are difficulties with the interpretation of current exemption to the requirement for consent for medical research, particularly where this involves the secondary use of data acquired from health services. Confusion remains about the nature of the obligation for researchers to notify participants that their data has been used (under the first data processing principle, namely that data has been fairly and lawfully processed).

36.2 Fair processing

The AMS Report 'Personal Data for public good' report analysed the requirements of the Data Protection Act and the requirements for fair processing. One of its conclusions was that the section 33 exemption for research was underutilised in part because regulators had interpreted a

need for additional fair processing requirements. Indeed its conclusion was as follows:

'Identifiable data can be used for medical research without consent, provided that such use is proportionate with respect to privacy and public interest benefits'.

There is a need for clarity as to the need for additional fair processing in population health research.

36.3 The scope of the obligation for lawful processing

Another area of debate is the extent to which the obligation for lawful processing necessarily connotes compliance with all other relevant legislation (including the common law of confidentiality and the Human Rights Act 1998) or whether a narrower interpretation which implies conformity with a condition in Schedule 2 or 3 (AMS report page 24). These questions of legal interpretation remain uncertain - and the AMS Report cited instances in which this ambiguity hampered epidemiological research.

PHG Foundation
6 October 2010

Contact details: <http://www.phgfoundation.org/contact/alison.hall>