Data Protection or exploitation?  
The erosion of safeguards for health and genetic research

May 2015

A new Data Protection Regulation is being negotiated by the European Union (EU), which will affect how personal information can be stored, used and shared in all EU member states.

This briefing describes proposed changes to the “research exemption” which, if adopted, would allow data to be stored and shared with commercial companies without people’s knowledge or consent. It considers the implications of these changes for the future storage and use of personal medical records and genetic data.

Summary

Proposed changes by the EU’s Council to data protection law include allowing:

- The further processing of health data, including genetic data, on a massive scale for research without people’s knowledge or consent;
- Indefinite retention of health data including genetic data such as whole genomes without people’s knowledge or consent;
- Sharing of this data with third parties, including companies such as Google, without people’s knowledge or consent, usually with names stripped off (pseudo-anonymised) but in a way which allows results to be fed back later to individuals later on, or other data sets (e.g. social care, education) to be linked in;
- Feedback of individual risk assessments calculated using such data through a change of use, even in cases where the individual never gave consent at the point of collection of the data for this to be analysed. Feedback could occur with “click on” consent and no requirement for medical involvement;
- Use of such risk assessments for marketing purposes, following individuals’ agreement to feedback, as these will be regarded as one of the “legitimate interests” of data controllers.

These are significant changes compared to existing law, which in general requires consent for research using health data which is not anonymised and also prevents data shared for research purposes without consent from being monetised through subsequent feedback of computer generated results to individuals (known as ‘profiling’). Allowing feedback of profiling from data processed without a person’s knowledge or consent creates a significant loophole as, in effect, all data can be shared for “research” purposes and stored indefinitely, whilst in reality being used for commercial purposes at a later stage. This is a back door route to health screening of whole populations by commercial companies without medical justification. Requirements to protect privacy, allow people a say over how their data is used and who has access to it, and prevent excessive government surveillance or commercial exploitation are all rendered meaningless by these proposals.

The proposals follow lobbying by the Wellcome Trust, which has proposed building a DNA database of all users of the National Health Service (NHS) in England as a public-private partnership in which Google and/or the Google-funded gene testing company 23andMe are expected to gain access to the data. The proposals are not compatible with:
• Ethical requirements for medical professionals, for example as enshrined in the Helsinki Declaration;
• Article 8 of the European Convention on Human Rights, which prevents the indefinite retention of biometrics such as genomes (“genetic fingerprints”) without an individual’s knowledge or consent, unless they have been convicted of a crime.

In contrast, the European Parliament’s amendments are compatible with ethical and human rights standards because they include a specific requirement for consent to research using health data, including genetic data.

Concerns about the Council’s leaked proposals include:
• Potential misuse for government and commercial surveillance through building a back door DNA database without people’s knowledge or consent;
• Commercial exploitation of research data through change of use to allow feedback of potentially misleading health risk assessments to individuals and their families and use of these for personalised marketing of medicines and other health products;
• Incompatibility with existing ethical and human rights standards, such as the Helsinki Declaration and the European Convention on Human Rights.

1. The ‘research exemption’ in existing law

The Data Protection Directive 95/46/EC contains key principles for the processing of personal data. This Directive allows EU Member States to make exceptions where justified in the public interest, including for “scientific research and government statistics”, provided governments “provide specific and suitable safeguards so as to protect the fundamental rights and the privacy of individuals”.

Under the current system, each member state interprets the requirements in the Directive into national law. As examples, in both the UK and Germany, personal data that is stored and analysed for research purposes can’t be used for other purposes.

In the UK, the relevant law is the Data Protection Act 1998, which states (Article 33) that data processed for research purposes must meet the conditions that: (a) that the data are not processed to support measures or decisions with respect to particular individuals, and (b) that the data are not processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject. In addition, the results of the research or any resulting statistics must not be made available in a form which identifies data subjects.

In Germany, the relevant law is the Federal Data Protection Act. Section 40 covers the processing and use of personal data by research institutes. This states that personal data collected or stored for scientific research purposes may be processed or used only for such purposes and may only be shared with non-public bodies if these undertake not to process or use the communicated data for other purposes. The personal data shall be de-personalized as soon as the research purpose permits this, and until then, identifying information must be stored separately and combined with the information only to the extent required by the research purpose. Bodies conducting scientific research may publish personal data only if the data subject has consented or this is indispensable for the presentation of research findings on contemporary events.

2. Proposed changes

In 2012, the European Commission proposed a major reform of the EU legal framework on the protection of personal data. It stated that the new proposals would strengthen individual
rights and tackle the challenges of globalisation and new technologies. The package of reforms include the development of a new General Data Protection Regulation, to cover most uses of personal data, and a new Data Protection Directive to cover uses of personal data for law enforcement and counter-terrorism.

The focus of this briefing is on the proposals for the new General Data Protection Regulation, which will replace the 1995 Data Protection Directive and cover most uses of personal data, including the use of health data and genetic data in research. A "directive" is a legislative act that sets out a goal that all EU countries must achieve. However, it is up to the individual countries to decide how. A "regulation" is a binding legislative act. It must be applied in its entirety across the EU.

The first draft General Data Protection Regulation was published by the Commission in January 2012. After scrutiny by the European Parliament’s Civil Liberties, Justice and Home Affairs Committee (the LIBE Committee), the Parliament adopted an amended version on 12th March 2014. The European Council (which represents the views of EU member state governments) then began negotiations to develop version of the Regulation. Once this has been finalised and published, the European Parliament will first consider it at the LIBE Committee. However, if the Council and Parliament do not immediately agree, preparations will begin for “trialogue” negotiations, most likely following the summer break. During trialogue, the European Commission, Parliament and Council agree a final version of the law, based on the three versions that they have developed. This joint text must then be approved by Parliament and Council before the Act can be adopted.

A leaked draft of the Council’s version of the draft Regulation, from December 2014, has been published and analysed online. A final version of the Council’s draft is expected in June 2015. Civil society groups were so concerned about the draft they wrote an open letter to the European Commission President asking him to keep the promise that data protection standards would not be weakened. However, the specific implications of changes to the “research exemption” have not yet been discussed.

2.1 The European Commission and Parliament’s draft Data Protection Regulations

The European Commission and European Parliament’s versions of the draft Regulation both make it clear that personal data does not include data that has been anonymised in such a way that the individual is no longer identifiable (Recital (23)): such data has always fallen outside the scope of data protection law. Both the Commission and Parliament require processing of personal data (i.e. potentially identifiable data) for health purposes or research purposes to meet particular conditions, laid out in Articles 81 and 83 respectively. Genetic data and health data are regarded as special categories of data (Article 9) and processing which is necessary for health or research purposes must also meet the conditions in Article 81 and/or Article 83, unless the data subject has given his or her consent. Fines for companies or organisations who don’t comply with the required conditions can be up to 500,000 Euros or 1% of the annual worldwide turnover of a commercial enterprise (Article 79(5)).

In the Commission’s view, personal data which is used solely for research may be stored for a longer period than necessary for other purposes, provided there is a periodic review of the necessity for storage (Article 5(e)): the European Parliament adds additional requirements to this Article to limit access to this data only for research purposes and provide additional safeguards. Under Article 17, the individual’s “right to erasure” of data can be overridden for reasons of public interest in the area of public health in accordance with Article 81; or for historical, statistical and scientific research purposes in accordance with Article 83.
As under the 1995 Directive, the proposed new Regulation allows specific exemptions to be made by law for processing sensitive personal data (including health and genetic data) for research purposes, provided suitable safeguards are adopted to protect personal data and fundamental rights (Recital (42)). The Commission’s text also gives powers to the Commission to adopt delegated acts covering research and other data uses (Recital (129) and Articles 9(3), 81(3), 83(3) and 86), rather than leaving these decisions solely in the hands of member states. Research is explicitly recognised to include privately-funded research (Recital (126)).

The European Parliament’s draft (but not the Commission’s one) foresees an exemption from the requirement of consent in cases of health research that serve a high public interest (Recital 123a). Importantly, it also notes that the processing of personal data for research purposes should not result in personal data being processed for other purposes, unless with the consent of the data subject or on the basis of Union or Member State law (Recital (126)).

The detailed conditions laid out in Articles 81 (health purposes) differ between the European Commission and Parliament’s draft versions of the text. Both versions require processing for health purposes to be necessary for medical purposes or management of healthcare services (in which case it must be done by a health professional or someone with an equivalent professional obligation of confidentiality), or for reasons of public interest in the area of public health (for which the Parliament also adds a requirement for a professional obligation of confidentiality), or for other reasons of public interest such as social protection and settling claims for benefits and health services in the health insurance system (where the Parliament adds a requirement that such data should not be used for other purposes without consent or a specific provision in law). The Commission requires processing of health data for research purposes to meet the conditions in Article 83, but the Parliament adds an additional requirement that such processing should only be permitted with the consent of the data subject (Article 81(2)), unless member state law has specifically provided for an exception (Article 81(2a)).

In Article 83, both versions allow research to be conducted using personal data without consent, provided the purpose can’t otherwise be fulfilled by processing data which no longer permits the identification of the data subject, and the identifying data is kept separately from other information. Such data is known as “pseudo-anonymised” because it may appear anonymised to the researcher using it, but the ability to link the data and its analysis back to the individual remains. However, the Parliament’s amendment in Article 81(2) means that, in its view, consent should normally be required for health research even if the conditions in Article 83 are met.

2.2 Changes made by the European Council

The European Council’s leaked draft of the Regulation refers to “processing for historical, statistical and scientific purposes”, rather than “research”. In contrast to the European Parliament’s view, the leaked Council draft states that further processing of personal data for statistical or scientific purposes should not be considered incompatible with the purposes for which the data are initially collected (Recital (125) and Article 1b) and that consent from the data subject should not be necessary for each further processing for scientific purposes (Recital (125aa)). Article 5(e) allows data to be stored for longer for statistical or scientific purposes, subject to “implementation of the appropriate technical and organisational measures”. In Article 6, the leaked Council draft retains the requirement that processing of personal data for statistical and scientific purposes must meet the conditions in (an altered version of) Article 83, but it removes the requirement that processing of health data should meet the requirements in Article 81 (which is deleted). Hence, the Council’s version omits the Parliament’s requirement for consent to use of health data for research in normal circumstances (with specific exceptions defined by law).
The Council’s new draft of Article 83 allows EU Member States to adopt laws which allow the processing of personal data for statistical or scientific purposes without consent. These laws may include exemptions from the rights to provide information to the data subject, access to the data, corrections of data, the right to erase and data portability and the right to object to processing. The new draft Article 83 refers to pseudo-anonymisation as a possible measure to safeguard data subjects’ interests. The leaked draft also states that if the result of scientific research in particular in the health context gives reason for further measures in the interest of the data subject, the general rules of this Regulation should apply in view of those measures (Recital 126). In contrast to existing law and to the European Parliament’s draft, this allows for a subsequent change of purpose (e.g. from research use to health use), provided this use meets regulatory requirements.

This is a significant change from existing law because data that has supposedly been shared for a limited purpose (i.e. for research) can then be used in ways which affect the individual. So, for example, if a company is given pseudo-anonymised medical records and genetic data for research – without an individual’s knowledge or consent - they can begin to analyse this data and develop computer algorithms to calculate the individual’s health risks or those of their family. Whereas existing law prevents these risks being fed back to individuals and used for other purposes (e.g. medical, insurance or marketing), the leaked Council’s draft Regulation allows these uses because “pseudo-anonymisation” means risk assessments can be linked back to the individual by the data controller later on. Re-identification and non-scientific use (such as for health purposes) would require consent to feedback of these risk assessments, but this is envisaged as requiring only a tick box on a website, perhaps via an individual’s online medical record. A major loophole in the draft law allows the “legitimate interests” of the data controller or a third party to justify further processing of data without consent: and this includes for marketing purposes. Thus, once an individual has agreed to feedback of risk assessments from a company who has their data, they can be marketed medicines, supplements or other products without consent, based on their predicted risks. This provides a means for companies to monetise the data, which does not exist under current data protection law.

The Council’s leaked draft also removes any reference to large fines for failing to implement required safeguards (Article 79).

In summary, the Council’s leaked draft implies a significant change from existing data protection law. It allows:

- Data to be used for research without consent when it has only been “pseudo-anonymised”, not anonymised. Pseudo-anonymisation means the results of the research can still be linked back to the individual by the data controller (who retains the identifying information), and other data sets with shared identifiers could also be linked in (e.g. social care, education, employment or tax records) without the individual’s knowledge or consent.
- Health data (e.g. from electronic medical records) and genetic data (including genotypes or whole genomes) to be stored indefinitely for statistical or scientific purposes and shared with commercial companies without people’s knowledge or consent (provided the data is pseudo-anonymised i.e. identifiers are stripped off but retained by the controller of the data so the data can later be linked back to the individual).
- Feedback of research results to individuals, via a change of purpose, even in cases where the individual never gave consent at the point of collection of the data for this to be analysed. This would allow computer calculations of health risks to be de-identified (linked back to individuals) and used for health, marketing or insurance purposes at a later stage. In short, it would allow commercial companies given
access to pseudo-anonymised data for research a means to subsequently monetise this data through a change of purpose.

3. Implications

There are two main areas of concern about the processing of health and genetic data:

- Human rights concerns because genotyping or genome sequencing is a dual-use technology which acts as a biometric identifier (a “genetic fingerprint”) for individuals and their families. This may allow police, governments, or others who may gain access to the databases, to track individuals and identify their relatives.

- Concerns that data-mining will lead to commercial exploitation of health and genetic data e.g. for marketing or insurance purposes, in ways that do not benefit, and may harm, the individual and their family, and undermine medical confidentiality, health services and public health.

3.1 Implications for human rights

If adopted without the safeguards added by the European Parliament, the new Data Protection Regulation would allow personal genetic information to be retained indefinitely in databases for research purposes, without people’s knowledge or consent. Although shared data would be pseudo-anonymised, this might not prevent people with access to these databases from deducing the identity of an individual, due to the vast information they could hold about them e.g. their age, where they live and what health conditions they have.

Further, because the data is pseudo-anonymised, there will be an institution (which may be government, charitable or private) which will retain the key and the ability to link this genetic information back to individuals’ names.

Whole genomes, or partial genotypes (based on millions of points along an individual’s DNA), act as a “genetic fingerprint” which can link all an individual’s personal data to their physical self and also identify their relatives. Thus, genomes or genotypes stored for medical or research purposes can be used in the same way as police DNA databases (which store a smaller portion of an individual’s genetic information, known as a DNA profile). Although such databases can be valuable in solving crimes, they can also be misused for surveillance purposes: for example, by tracking down who attended a political meeting by collecting the DNA left in their saliva on their coffee cups. The creation of large-scale databases allows this use because the genomes or genotypes are searchable and linked to people’s names and addresses, so that the person who has left their DNA in a particular place can be identified.

Unlike fingerprints, genetic fingerprints can also be used to identify a person’s biological relatives, including their children.

In future, newly developed rapid DNA testing systems, which analyse saliva samples to produce genetic data on the spot, could be used at borders, or even on the street, to match an individual to their online records and potentially not only track their whereabouts but also retrieve linked medical records and other data (such as social care and education records) and those of their children or other close relatives.

Police and security services have the powers to access health data, including genetic data, if it is being stored. In the USA, there has already been one recent case of police access to a commercial genetic database leading to an innocent man being caught up in a 20-year-old murder investigation. If everyone’s genome is stored on a database, there is little doubt that police and security services would be able to use this to track individuals and their relatives. As the Edward Snowden revelations showed, this could include “back door” access to genomes stored in the cloud by national or foreign security services. Although these
powers might be used for solving crimes, they could also allow repressive regimes to track down dissidents or identify their relatives.

The main safeguard to prevent excessive government or police surveillance is the right to erasure of genetic information and fingerprints (known as “biometrics”) from databases, unless retention of such data is necessary for the prevention of disorder or crime (Article 8 of the European Convention of Human Rights). This right was confirmed in the 2008 judgment of the Grand Chamber of the European Court of Human Rights, which led to the removal of innocent people’s DNA profiles from the UK police National DNA Database. However, the Council’s leaked draft of the Regulation removes any right for an individual to seek removal of their, or their child’s, genetic data from government or third party (including commercial) databases, as it allows this data to be retained indefinitely for research purposes, as well as shared with third parties without the individual’s knowledge or consent. As well as breaching the European Convention on Human Rights, this Council proposal is likely to breach the EU Charter of Fundamental Rights (Articles 3, 6, 7 and 8), since it is hard to see how provisions which allow Member States to build a DNA database of the population without consent, with no right of removal, can be considered compatible with fundamental human rights.

### 3.2 Commercial exploitation and health implications

If adopted without the safeguards added by the European Parliament, the new Data Protection Regulation would allow electronic medical records to be handed over en masse to one or more commercial companies without individuals' knowledge or consent, provided the data was pseudo-anonymised. Where genetic data has been collected (or is able to be collected in the future for any purpose) this could also be transferred to commercial companies without people's knowledge or consent.

These proposals are incompatible with ethical standards which require people to be asked for their consent to take part in health research after being informed, for example, about potential commercial conflicts-of-interest. For example, the Helsinki Declaration is a set of principles for medical professionals conducting research. The Helsinki Declaration includes requirements to protect the "dignity, integrity, right to self-determination, privacy and confidentiality of personal information of research subjects". Research participants must be informed of "the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliations of the researcher, the anticipated benefits and potential risks of the study and the discomfort it may entail, and any other relevant aspects of the study" before giving their consent, preferably in writing. For consent to valid it must be fully informed and freely given. Special protections must be accorded to people who lack capacity to give consent and account must be taken of the changing capacity of children as they grow up. The European Convention on Biomedicine states (Article 16): Research on a person may only be undertaken if all the following conditions are met:

(i) there is no alternative of comparable effectiveness to research on humans;
(ii) the risks which may be incurred by that person are not disproportionate to the potential benefits of the research;
(iii) the research project has been approved by the competent body after independent examination of its scientific merit, including assessment of the importance of the aim of the research, and multidisciplinary review of its ethical acceptability;
(iv) the persons undergoing research have been informed of their rights and the safeguards prescribed by law for their protection;
(v) the necessary consent as provided for under Article 5 has been given expressly, specifically and is documented. Such consent may be freely withdrawn at any time.

Unless the European Parliament’s proposed safeguards are included, the rights of patients to be informed about who has access to their personal data and how it will be used will have been removed. Allowing such a major role for commercial interests risks a major loss of
public trust in medical research and patient confidentiality. The medical profession, who collect blood samples and other health information, will also have been compromised.

Development of the Data Protection Regulation has been led the Commission’s Justice Directorate and member states’ justice departments, not by health departments. Yet it could have a profound impact on health services and on trust in medical confidentiality throughout member states.

The UK-based research charity, the Wellcome Trust is the main organisation that has been lobbying to remove requirements for consent from the new Data Protection Regulation. In the UK, the Wellcome Trust (which played a major role in the Human Genome Project) and the UK Government’s Human Genome Strategy Group have proposed building a DNA database of the whole population by including whole genomes as attachments to electronic medical records in the National Health Service (NHS). The UK Government has backed this plan and the Secretary of State for Health, who remains in post after the Conservative victory in the 2015 election, has stated that in his view every baby should have its whole genome sequenced at birth. Such a project would be extremely costly and could only take place as a public-private partnership, with significant commercial control over the research agenda. This database cannot be built with consent from every member of the public because most people would not sign up to it: for example, the recruitment rate for the UK Biobank research project was only 5 to 10% of the people it approached.

The first stage of the plan to build this database, known as “care.data” has been highly controversial, because it involves building a central database of electronic medical records in the NHS, without seeking consent, and sharing this information with private companies. Due to significant public controversy, “Care.data” was put on hold until after the May 2015 UK election. However, the Life Sciences Minister George Freeman informed the Daily Mail newspaper that a Conservative Government would revive plans to give private companies such as Google responsibility for storing people’s private personal health data.

Google, and the Google-funded gene test company 23andMe, both have a major interest in obtaining access to health and genetic data.

Google has been in negotiations with the UK Department of Health about access to NHS medical records and genomes since at least 2008. The Wellcome Trust and UK Medical Research Council have already funded joint research with 23andMe, so there is little doubt that Google – which is planning a major to open a major new HQ in London in 2017 – will be given access to this data, provided there is no requirement to ask individuals for their consent (which many may be expected to refuse). In 2013, Patrick Chung, a 23andMe board member and partner at the venture-capital firm NEA 23andMe told Fast Company: “...23andMe will make money by partnering with countries that rely on a single-payer health system. "Let's say you genotype everyone in Canada or the United Kingdom or Abu Dhabi," he says, "and the government is able to identify those segments of the population that are most at risk for heart disease or breast cancer or Parkinson's. You can target them with preventative messages, make sure they're examined more frequently, and in the end live healthier lives, and the government will save massive expenses because they halted someone who's prediabetic from getting diabetes. 23andMe has been in discussion with a bunch of such societies".

However, 23andMe’s investor, Google, is an advertising company, and it is not hard to see that in order to make money from “preventative messages” it must use personalised marketing. The use of health risk assessments, including genetic risk assessments, for marketing is a major driver of commercial interest in this area. Given access to people’s genomes and health data, there is enormous potential for this data to be exploited for personalised marketing by selling individuals, or their relatives, medicines, supplements and
foods which supposedly reduce their calculated risks. Fifteen years ago, the former chair of pharmaceutical company GlaxoSmithKline advocated retaining the NHS only as a basic service whilst using genetic screening as a means to personalise medical care and massively expand “pre-symptomatic” treatment (and pharmaceutical company profits). In fact, the Mail has already exposed how some private medical information from NHS patients is being sold and used for marketing. Mass-marketing of health products to healthy people is likely to lead to unnecessary tests and treatments which could harm people’s health.

Importantly, there is no evidence that such targeted “preventative messages” or increased use of medicines or supplements by healthy people will be good for health. Genes are in reality poor predictors of most diseases in most people and genetic tests are in general not useful as a way to decide who should stop smoking, eat healthily, exercise or take particular medicines or supplements. In the USA, the Food and Drug Administration (FDA) banned 23andMe from selling genetic test results in 2013, as the company was unable to supply evidence that its algorithms for calculating people’s health risks were scientifically valid. Gene tests are currently not regulated in the EU, allowing 23andMe to begin a sales push in Europe, including marketing its tests on British High Streets in Superdrug. A draft EU regulation which could change this situation (the In Vitro Diagnostics, or IVD, Regulation) has yet to be adopted. There is therefore no mechanism to prevent misleading claims about health risks from being marketed.

Conclusions

Proposed changes to EU Data Protection Regulation which allow personal health and genetic data to be stored indefinitely and shared with private companies without the knowledge or consent of individuals are incompatible with human rights, because:

(i) They allow genomes to be retained indefinitely, building a back door DNA database of whole populations without people’s knowledge or consent; and

(ii) They allow commercial exploitation of personal data through later feedback of risk assessments to individuals, which can then be used for marketing or other purposes. This is a back door route to screening of whole populations by commercial companies without medical justification, and without people’s knowledge or consent until the final feedback stage (where online click-on consent may be all that is required). There is currently no regulation of these commercial health risk assessments, which may mislead people about their health and lead to over-treatment.

Such legislation would be in breach of the European Convention on Human Rights (Article 8, right to privacy) and the Helsinki Declaration, which requires medical professionals to inform their patients of how their data will be used, including any conflicts-of-interest.

Pseudo-anonymised data (i.e. data with identifiers such as names stripped off, but which can still be linked to other data sets and to the individual) is still personal data, otherwise it would not need to be covered by data protection laws. Indefinite storage and widespread sharing of medical and genetic data could lead to excessive government or corporate surveillance of individuals and their families and to commercial exploitation through personalised marketing.
References


13 See Article 6 and Recital (39); and Data Protection: Broken Badly. EDRI, Access, Privacy International and Fundacja Panoptykon. https://edri.org/files/DP_BrokenBadly.pdf


GeneWatch UK Briefing
May 2015
Children could have DNA tested at birth. The Telegraph. 8th December 2013. http://www.telegraph.co.uk/health/healthnews/10501788/Children-could-have-DNA-tested-at-birth.html


