Article

MEDICAL LAW INTERNATIONAL

Medical Law International
2014, Vol. 14(4) 187–212

DOI: 10.1177/0968533215571956
mli.sagepub.com
SAGE

A comparative analysis of the requirements for the use of data in biobanks based in Finland, Germany, the Netherlands, Norway and the United Kingdom

Linda Briceño Moraia and Jane Kaye
University of Oxford, UK

Anne Marie Tasse and Bartha M. Knoppers
McGill University, Canada

Colin Mitchell
University of Oxford, UK

Sirpa Soini
National Institute for Health and Welfare (THL), Finland

Nils Hoppe
Leibniz University of Hanover, Germany

Susan E. Wallace
University of Leicester, UK

Morten Øien
Norwegian University of Science and Technology (NTNU), Norway

Corresponding author:
Linda Briceño Moraia, Centre for Health, Law and Emerging Technologies (HeLEX), Nuffield Department of Population Health, University of Oxford, Old Road Campus, Headington, Oxford, OX3 7LF, UK.
Email: linda.bricenomoraia@dph.ox.ac.uk

Abstract
To understand the causes of disease and improve diagnosis and treatment regimes, biomedical researchers need access to large numbers of well-characterized data and samples. Over the past decade, biobanks have been established across Europe to collect and manage access to data and samples. The challenge that we face is how to develop the tools and collaborations to enable researchers to access samples and data from a network of biobanks, rather than applying to individual biobanks. One of the perceived stumbling blocks to achieving this is represented by the different legal requirements in each country. The aim of the BioSHaRE-European Union (EU) project is to address these challenges by developing tools and methods for researchers to access and use pooled data from different cohort and biobank studies. The purpose of this article is to identify and compare the key legal requirements regarding research use of data across biobanks based in Finland, Germany, the Netherlands, Norway and the UK. Our investigation starts with the analysis of the key differences for the use of data between these countries. As a result, we identified three key areas where legal requirements differ across the five BioSHaRE-EU jurisdictions, namely, in the definition of personal data, the requirements regarding pseudonymization and processing for medical research purposes. This article provides an overview of these differences and describes them in the light of the proposed EU regulation on data protection.

Keywords
Access to data, biobanks, biomedical research, cross-border, data protection

Introduction

The role of biobanks

Biomedical research relies on the collection of personal information and human tissues held in biobanks,[1] in order to obtain the sample sizes needed to understand the aetiology of a given disease. Research is increasingly international, and samples obtained from many different biobanks, located in different jurisdictions, are now gathered to facilitate the creation of mega data sets.[2]

1. On a legislative level, there is no uniform definition of a biobank. Furthermore, it is not clear whether the term 'biobank' includes samples and related data. For the purpose of this article, the term biobank is defined as 'an organized collection of human biological material and associated information stored for one or more research purposes'. Available at: http://www.p3g.org/biobank-lexicon (accessed 8 April 2014).
2. G. Church, C. Heeney, N. Hawkins, J. de Vries, P. Boddington, J. Kaye, M. Bobrow and B. Weir, 'Public Access to Genome-wide Data: Five Views on Balancing Research with Privacy and Protection', *PLoS Genet* 5(10) (2009), p. 1; J. Gillot, 'Human rights, privacy and medical research: analysing UK Policy on tissue and data', *Genetic Interest Group* (2006). Available at: http://www.bionews.org.uk/page_37875.asp; G. Laurie, P. Mallia, D. A. Frenkel, A. Krajewska, H. Moniz, S. Nordal, C. Pitz and J. Sandor, 'Managing Access

Biobanks are fundamental research tools for the advancement of biomedical research. Their value is not limited to the stored samples and data, but goes beyond, and includes all the information that can be obtained from the analysis of these samples, the possibility to combine it with other relevant data, including clinical data, and the fact that they enable long-term research projects providing access to those resources over time. This is particularly true for large-scale longitudinal biobanks, which are generally used for epidemiological research projects and population biobanks, which usually store sensitive health data of thousands of individuals and may provide the basis for research in the field of personalized medicine. Small collections of samples and data, in academic or hospital settings, are also fundamental to contribute in the advancement of specific research projects, particularly when combined with results deriving from different studies. In general, biobanks may differ under different aspects, such as size, research design and types of biological samples collected, method of sample/data collection, processing, storage, and the research focus. The same diversity can be found in the relevant regulatory framework, but such jurisdictional differences may represent an obstacle for the pooling of data across the different biobanks or cohorts.

A fragmented regulatory framework

The law that applies to biomedical research within Europe is complex. It brings in different heads of law, such as data protection, privacy and tissue regulation as well as medical research regulations. At the core of this multifaceted legal framework is a distinction between sample and data that are governed by different legal regimens.[3] Whilst there is not one European legal instrument that applies specifically to biobanks, there are specific national biobanking legislations in some jurisdictions. One of the concerns is that although researchers operate across borders, the regulatory framework for biobanks is nationally based, and this may impede the pooling of data from biobanks.[4]

The laws relating to the processing of personal data in Europe share a common origin in the Data Protection Directive 95/46/EC (in this article simply 'the Directive'),[5] but there are some differences between member states regarding its implementation, particularly, in regard to biomedical research, which is most evident in the interpretation of key principles. The purpose of this article is to identify and compare the key legal requirements regarding research use of data in Finland, Germany, the Netherlands, Norway and the UK. All these countries are part of the BioSHaRE-European Union (EU)

---

to Biobanks: How Can We Reconcile Individual Privacy and Public', *Medical Law International* 10 (2010), p. 315–337.

3. T. Schulte in den Bäumen et al., 'Data Protection and Sample Management in Biobanking – A Legal Dichotomy', *Genomics, Society and Policy* 6 (2010), pp. 33–46.
4. J. Bovenberg, 'Legal Pathways for Cross-border Research: Building a Legal Platform for Biomedical Academia', *European Journal of Human Genetics* 15 (2007), pp. 522–524.
5. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L 281. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML (accessed 8 July 2014).

project,[6] which aims to develop tools to enable statistical analyses of large-scale data across different studies. The question underlying our analysis is whether the different requirements for accessing data in each of these jurisdictions may make it difficult for researchers to conduct studies that use samples and data obtained from a number of biobanks. We discovered firstly that there are considerable differences in the way that the data protection principles are implemented into national law, which reflect the different legal traditions of the BioSHaRE-EU countries. Secondly, there are three key areas of considerable differences between the requirements in these countries. These are the definition of 'personal data', requirements for pseudonymization and rules on processing of data for medical research, and they do not appear to have been addressed in the current version of the proposed EU Data Protection Regulation.

Those differences imply that researchers seeking access to biobanks located in different countries may be required to be compliant with different rules. For example, such a fragmented framework creates bureaucratic impediments for researchers,[7] as well as an increasing sense of uncertainty regarding the legal requirements found in the different EU member states, discouraging collaboration between researchers and, therefore, innovation. Since there are different types of biobanks, it seems that there is a need for a specific and unique body of regulations at European or international level, which could better address the legal challenges raised by biobanking research than the current general and fragmented system.

Data protection law

Legal framework across BioSHaRE-EU jurisdictions

Biobanks contain samples and associated clinical data, so the law that applies to the use of data for medical research purposes is highly relevant to biobanking activities. The laws relating to the processing of personal data in Europe share a common origin in the Data Protection Directive,[8] and therefore member states should implement the same core principles regarding the processing, storage and uses of data. The Article 29 Data Protection Working Party[9] set up under this Directive has been highly influential in the interpretation of the Directive's principles and provisions. The Working Party is composed of representatives from the member states' Data Protection Authorities, the EU Commission and the EU Data Protection Supervisor, which is an independent authority.

6. BioSHaRE is a consortium of leading biobanks and international researchers from all domains of biobanking science. Available at: https://www.bioshare.eu/ (accessed 8 April 2014).
7. F. Colledge, B. Elger and H.C. Howard, 'A Review of the Barriers to Sharing in Biobanking', *Biopreservation and Biobanking* 11(6) (2013), p. 343.
8. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
9. Available at: http://ec.europa.eu/justice/data-protection/article-29/index_en.htm (accessed 8 April 2014).

The margin of appreciation that member states can exercise in the implementation of Directives is particularly evident in regard to the Data Protection Directive.[10] In our case study of the BioSHaRE-EU jurisdictions, there are differences in the way that the Directive is implemented, which reflect the jurisprudential traditions of member states and whether they have specific legislation that applies to patients or medical research. The common approach is to develop national data protection acts based on the EU Data Protection Directive. However, this legislation does not always contain all the relevant rules. Instead, some are found elsewhere in other national legal instruments. Most countries also have a specific legislation on medical research that must be read along with the data protection legislation to provide a comprehensive picture of the legal requirements that apply to medical research.

In Finland, the Personal Data Act (523/1999) implemented the Data Protection Directive and governs research use of data.[11] There are also specific laws that apply to patients, which give strength to the data protection provisions, such as the Patient Rights Act (785/1992). As in other jurisdictions, the Personal Data Act is *lex generalis.* This means that it provides an overall framework for the use of data, and there are specific provisions that apply to medical research elsewhere. This is common to all the other BioSHaRE-EU jurisdictions.

In Germany, rather than detailed procedural legal provisions regulating the use of tissues and data, biomedical research relies on respecting constitutionally enshrined individual rights, with strong jurisprudential links to dignity and personality rights.[12] Accompanying these national rights, as in all of the BioSHaRE-EU jurisdictions, are the

10. See in general also J.-E. Litton and J. Bovenberg, 'To Explore Pan-European Solutions for the Cross-border Data Protection Issues Associated with BBMRI', *BBMRI Joint Deliverable WP5* (2011); T. Lemmens and L. Austin, 'The End of Individual Control over Health Information: Promoting Fair Information Practices and Governance of Biobank Research', in J. Kaye and M. Stranger, eds. *Principles and Practice in Biobank Governance.* (Farnham: Ashgate, 2009), p. 243.

11. Section 8 states that personal data shall be processed only if the data subject (i) has unambiguously consented to it, (ii) has given an assignment for the same or (iii) processing is necessary (inter alia) in order to (a) perform a contract to which the data subject is a party; (b) take steps at the request of the data subject before entering into a contract; (c) protect the vital interests of the data subject (in an individual case); (d) processing is based on the provisions of an Act or it is necessary for compliance with a task or obligation to which the controller is bound by virtue of an Act or an order issued on the basis of an Act; (e) there is a relevant connection between the data subject and the operations of the controller, based on the data subject being a client or member of, or in the service of, the controller or on a comparable relationship between the two (connection requirement). In these cases personal data may be disclosed only if (i) such disclosure is a regular feature of the operations concerned, (ii) the purpose for which the data is disclosed is not incompatible with the purposes of the processing and (iii) it can be assumed that the data subject is aware of such disclosure. English translation is available at: http://www.finlex.fi/fi/laki/kaannokset/1999/en19990523.pdf (accessed 8 April 2014).

12. J. Simon et al, 'A Legal Framework for Biobanking: The German Experience', *European Journal of Human Genetics* 15 (2007), pp. 528–532.

standard, harmonized data protection provisions present through the implementation of the Data Protection Directive (with minor deviations, as shown below) through the Federal Data Protection Act (DPA) (Bundesdatenschutzgesetz),[13] which applies only to the bodies mentioned in section 1(2).[14] For other bodies, the DPAs of the Länder (federal states) are applicable. Additional norms are found in canonical law (for hospitals run by the churches) and each federal state has a variety of differing data protection and hospital laws that address data protection issues, that is, the applicability of these norms depends on the legal status of the entity using the data. The fragmentation of the sources of law in this country is therefore considerably higher when compared to the other BioSHaRE-EU countries.

In the Netherlands, in addition to the Personal Data Protection Act (PDPA) of 2000 (*Wet bescherming persoonsgegevens*), which implements the European Commission (EC) Directive, there are other laws, not specifically addressing biobanks but specifically addressing the collection and linking of data for (biomedical) research (FEDERA codes).[15] The Dutch Civil Code (Burgerlijk Wetboek) also contains provisions that relate to the transfer of personal information from medical records for research purposes, without the need for consent, which are similar to the ones implemented in the PDPA. The two provisions differ slightly and in case of a conflict of terms, the more stringent provisions from Civil Code prevail.

In contrast to the other BioSHaRE-EU jurisdictions, the Norwegian legislative framework puts all research use of personal health data and samples under the same legislative instrument, known as the Health Research Act 2008 (HRA; Helseforskningsloven).[16]

13. Federal Data Protection Act (BDSG), in the version promulgated on 14 January 2003 (Federal Law Gazette I, p. 66), last amended by Article 1 of the Act of 14 August 2009 (Federal Law Gazette I, p. 2814), in force from 1 September 2009. An English version of the Act is available at: http://www.bfdi.bund.de/EN/DataProtectionActs/Artikel/BDSG_idFv01092009.pdf?__blob=publicationFile (accessed 8 July 2014).
14. That is (1) public bodies of the Federation; (2) public bodies of the Länder insofar as data protection is not governed by Land legislation and insofar as they (a) execute federal law or (b) act as bodies of the judicature and are not dealing with administrative matters; (3) private bodies insofar as they process or use data by means of data processing systems or collect data for such systems, process or use data in or from non-automated filing systems or collect data for such systems, except where the collection, processing or use of such data is effected solely for personal or family activities. Federal Data Protection Act (BDSG), section 1(2).
15. See the Codes of Conduct for medical research set out by the Dutch Council of the Federation of Medical Scientific Societies (*Federatie Van Medisch Wetenschappelijke Verenigingen*). Available at: http://www.federa.org/gedragscodes-codes-conduct-en (accessed 8 April 2014).
16. The Health Research Act 2008 (Helseforskningsloven) has superseded the Personal Data Act 2000, as well as Patients' Rights Act (1999 No. 63) and the Personal Health Data Filing System Act 2001 (Act of May 2001 No. 24 on Personal Health Data Filing Systems and the Processing of Personal Health Data), Biobanks Act 2003, Application of Biotechnology in Human Medicine Health Act 2003.This Act applies only to research conducted in Norway or conducted under the direction of a person or body responsible

It applies to all medical and health research, which is defined as 'the use of scientific methods to generate new knowledge about health and disease', and it applies to research conducted on human beings, human biological material or personal health data, including pilot studies and experimental treatments.[17] There are also common rules regarding the use of human biological material and personal health data.

In the UK, there are no biobank-specific laws but rather a combination of general statutory and regulatory provisions, common law doctrines and guidance documents.[18] Rules governing access to data are contained in the DPA 1998 and in the Data Protection (Processing of Sensitive Personal Data) Order 2000,[19] making it relatively easy to find out where the Directive provisions have been implemented. Unlike Finland, Norway and the Netherlands where there is specific legislation that applies to medical research, the UK has no specific statute for patient rights and medical research. The closest that the UK gets in this regard are the Medicines for Human Use (Clinical Trials) Regulations 2004, which set out good practice in the conduct of clinical trials.[20]

The Data Protection Directive is to be replaced with a new regulation,[21] which, unlike Directives, would have direct effect in all EU member states without having to be implemented in national law. The intention is that the new Data Protection Regulation will result in a single set of rules across Europe. However, although expressly intended to strengthen the rights of individuals, current drafts of the Regulation could potentially have a negative impact on health research.[22] There has been concern as to whether the

---

for the research established in Norway. It does not apply to use of personal health data if (i) the person or body responsible for the research is established in another European Economic Area (EEA) State; (ii) the person or body responsible for the research is established in a state outside the EEA and the institution does not use tools in Norway for purposes other than pure transfer of personal health data. Clinical testing of medicinal products on human beings is covered by the Medicines Act and appurtenant regulations. Clinical testing of medical equipment is covered by the Act on Medical Equipment and appurtenant regulations. In this context, the provisions of the Biobank Act apply as a supplement, where relevant.

17. See para. 2 of the Health Research Act.
18. L. Curren, P. Boddington, H. Gowans, N. Hawkins, N. Kanellopoulou, J. Kaye J, and K. Melham, 'Identifiability, Genomics and U.K. Data Protection Law', *European Journal of Health Law* 17(4) (2010), 329–344; S.M.C. Gibbons, *Elsagen UK Law Report*, (April 2004), 14.
19. Data Protection Act 1998, 16 July 1998, Available at: http://www.legislation.gov.uk/ukpga/1998/29/contents and The Data Protection (Processing of Sensitive Personal Data) Order 2000 (accessed 17 February 2000). Available at: http://www.legislation.gov.uk/uksi/2000/417/introduction/made (accessed 8 April 2014).
20. Medicines for Human Use (Clinical Trials) Regulations 2004 (SI 2004/ 1031)
21. The latest draft is available at: http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P7-TA-2014-0212&language=EN&ring=A7-2013-0402 (accessed 8 July 2014).
22. See the 'Wellcome Trust open letter re: amendments to EU Data Protection Regulation', Available at: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/WTP055585.pdf (accessed 8 April 2014), and the

proposed regulation will prohibit the use of a broad consent in biobanking and much of the epidemiological research carried out on registries.[23]

Key differences

Just as significant variation exists in the way that the BioSHaRE-EU countries have implemented and enforced the provisions of the Data Protection Directive, there are also differences in the way that the various elements of the Directive have been interpreted. This makes for considerable complexity and requires specialist knowledge of local interpretation. There are three key areas of difference in interpretation, namely, the definition of personal data, the notion of pseudonymization and the requirements for processing of data for research purposes.

Personal data. Biobanks provide access not only to human samples but also to different medical information, such as genomic data, population genetic data and molecular data. In this framework, data sharing may refer to the transfer of any type of personal and/or sensitive data referring to the donor, clinical data as well as data deriving from the analysis of samples.[24]

The definition of personal and sensitive data, as mentioned in the EU Directive, has been implemented by member states. It has a broad scope and therefore leaves space to a broad interpretation. Even if there are some slight differences in how it has been

---

'Impact of the draft European Data Protection Regulation and proposed amendments from the rapporteur of the LIBE committee on scientific research', available at: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtvm054713.pdf (accessed 8 April 2014).

23. Jan Philipp Albrecht; 'rapporteur' for the data protection dossier:

    DRAFT REPORT on the proposal for a regulation of the European Parliament and of the Council on the protection of individual with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) (COM(2012)0011 – C7-0025/2012 – 2012/0011(COD)).

    Available at: http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/pr/922/922387/922387en.pdf (accessed 8 July 2014). On the possible impact of the Draft Regulation (before the Albrecht amendments), see G. Lauss, A. Bialobrzeski, M. Korkhaus, K. Snell, J. Starkbaum, A.E. Vermeer, J. Weigel, H. Gottweis, I. Helén, J. Taupitz and P. Dabrock, 'Beyond genetic privacy. Past, Present and Future of Bioinformation Control Regimes', (2013), 45. Available at: http://private-gen.eu/uploads/media/PRIVATE_Gen_FINAL-REPORT_2013_01.pdf (accessed 2 February 2015). On the impact of the EU Regulation on biomedical research after Albrecht amendments, see M.C. Ploem, M.L. Essink-Bot, 'Proposed EU Data Protection Regulation is a Threat to Medical Research', *British Medical Journal* 364 (2013), f3534; Protecting health and scientific research in the Data Protection Regulation (2012/0011(COD)): Position of non-commercial research organisations and academics - July 2014. Available at: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/WTP055584.pdf (accessed 2 February 2014); L. Briceno Moraia and J. Kaye, 'Spies, Data and Research', *EMBO Reports*, 15(2) (2014), p. 123.

24. F. Colledge, B. Elger and H.C. Howard, 'A Review of the Barriers to Sharing in Biobanking', *Biopreservation and Biobanking* 11(6) (2013), p. 344.

implemented by BioSHaRE-EU jurisdictions, they are not likely to have an impact on practice. Nevertheless, different interpretations of what data could be considered as personal data, and therefore of what enables identification of the individual, may influence the assessment of which data, under certain conditions, shall be anonymized or pseudonomized (see subsection (b) below).

Under Article 2(a) of the Data Protection Directive:

> "personal data" shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.

There are two important elements to this definition. Firstly, the data subject must be a natural person, and secondly, the list of factors that can make an individual identifiable. The Directive does not provide any other detailed rules with regard to genetic data, but the latest draft of the EU Data Protection Regulation includes this kind of data, following European Court of Human Rights case law.[25] Prior to this, the Article 29 Working Party introduced the topic of the protection of genetic data in its Work Programme in 2003 and highlighted the fact that even though some member states explicitly listed genetic data as sensitive data, this processing is not always regulated by specific legislation but by complementary rules on patient's rights.[26] It is in this area where there are specific differences between five of the BioSHaRE-EU countries.

Each of the BioSHaRE-EU jurisdictions adhere to the spirit of the Directive definition, by stating that personal data are data that relate to an identified or identifiable natural person. However, there are subtle differences between the BioSHaRE-EU jurisdictions on even this most basic element of the Directive. For example, under Article 1(a) of the Dutch Data Protection Directive, 'personal data' means any information relating to an identified or identifiable natural person, but under section 3(1) of the German Data Protection Act personal data means 'any information concerning the personal or material circumstances of an identified or identifiable individual (the data subject)'. 'Material circumstances' is much broader than the comprehensive list of items that makes up the Directive's definition under Article 2(a) and a broader approach may have been endorsed by the announcement in the Work Programme 2007 that there should be avoidance of 'unduly restricting the interpretation of the concept of personal data'.[27]

25. See Article 9 of the Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (COM(2012)0011 Available at: http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2F%2FEP%2F%2FTEXT%2BREPORT%2BA7-2013-0402%2B0%2BDOC%2BXML%2BV0%2F%2FEN&language=EN#title1. (accessed 8 July 2014).
26. See Article 29 Data Protection Working Party, 12178/03/EN WP 91, Working Document on Genetic Data, 17 March 2004, 3. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2004/wp91_en.pdf (accessed 8 April 2014).
27. Opinion 4/2007 on the concept of personal data, WP136 (2007), Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf

Whilst the Directive refers in general to a 'natural person', section 1(1) of the UK DPA 1998 defines personal data as data relating to a 'living individual'. The living individual is someone who can be identified (a) from those data or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller and include any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual. As with other BioSHaRE-EU countries, this definition of personal data recalls the text of Article 2 of the EC Directive but with the peculiarity that the UK DPA specifies the variety of sources from which the information may be derived[28] rather than the factors to which it may be related (physical, physiological, mental etc.). The inclusion of any data or information *likely to* come into the possession of the data controller means that this definition of personal data is broader than in the other BioSHaRE-EU countries. The UK DPA definition of 'sensitive personal data' recalls the categories to which the Directive refers and contains a list of kinds of information that can be considered sensitive data, including any data relating to the physical or mental health condition of the data subject (section 2(e)). This is another difference in legal approach because even though health data are implicitly considered sensitive personal data in all the BioSHaRE-EU jurisdictions, only in the UK and in Norway is there an explicit reference to physical or mental health data in law.

It is in Finland and in Norway where there is greatest deviation from the specifications of the Data Protection Directive. Whilst the Finnish DPA basically follows the requirements of the Directive, there is a difference in the interpretation of personal data. Under the Finnish DPA, personal data means 'any information on a private individual and any information on his/her personal characteristics or personal circumstances, where these are identifiable as concerning him/her or the members of his/her family or household' (section 3).[29] This is a significant difference because it includes information relating to 'members of his/her family or household' and is unlike any of the other jurisdictions. Arguably, however, such an approach is in line with the Article 29 Working Group's inclusion of genetic information as 'personal information' and also with the intent of the proposed EU Data Protection Regulation to support a broad definition of personal data.

---

28. Under Article 1(1) of the UK Data Protection Act:

> unless the context otherwise requires "data" means information which – (a) is being processed by means of equipment operating automatically in response to instructions given for that purpose, (b) is recorded with the intention that it should be processed by means of such equipment, (c) is recorded as part of a relevant filing system or with the intention that it should form part of a relevant filing system, [F1 . . .] (d) does not fall within paragraph (a), (b) or (c) but forms part of an accessible record as defined by section 68; [F2] or (e) is recorded information held by a public authority and does not fall within any of paragraphs (a) to (d); [. . .].

29. And in particular, 'if they relate to or are intended to relate to [. . .] the state of health, illness or handicap of a person or the treatment or other comparable measures directed at the person' (section 11, (4)); (2) processing of personal data means 'the collection, recording, organisation, use, transfer, disclosure, storage, manipulation, combination, protection, deletion and erasure of personal data, as well as other measures directed at personal data'.

In Norway, under para. 4(d) of the Norwegian HRA, personal health data are defined as 'confidential information pursuant to Section 21 of the Health Personnel Act and other information and assessments concerning health issues or that are significant for health issues that can be linked to an individual person'. This definition is broad enough to include genetic data as suggested by the Article 29 Working Party and in the proposed draft of the Data Protection Regulation under Article 4(2) and (10) (definitions).[30] Article 4(10) further specifies that:

> ... "genetic data" means all personal data relating to the genetic characteristics of an individual which have been inherited or acquired as they result from an analysis of a biological sample from the individual in question, in particular by chromosomal, deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) analysis or analysis of any other element enabling equivalent information to be obtained.

Pseudonymization. Anonymization and consent are used to protect privacy interests, but both may restrict data sharing. In some cases, such as in the field of biobanking research, it is recognized that it is not possible to reach a complete anonymization[31] and therefore the threshold is limited to the pseudonymization requirement, and consent is not required if certain conditions, including coding, are fulfilled.

Another area of national variation in the interpretation of the EU Directive is the definition of pseudonymization. Under the Directive, data protection principles shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable. To determine whether a person is identifiable 'account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person' (Recital 26).[32] The Directive does not provide further rules regarding anonymized, coded or pseudonymized

30. 'Personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, unique identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social or gender identity of that person, Available at: http://www.janalbrecht.eu/fileadmin/material/Dokumente/DPR-Regulation-inofficial-consolidated-LIBE.pdf (accessed 8 April 2014).
31. J. Kaye, N. Kanellopoulou, N. Hawkins, H. Gowans, L. Curren and K. Melham, 'Can I Access my Personal Genome? The Current Legal Position in the UK', *Medical Law Review* 22(1) (2014), p. 64; N. Homer and others, 'Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures using High-density SNP Genotyping Microarrays', *PLoS Genet* 4(8) (2008), p. e1000167. Available at: http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1000167 (accessed 24 October 2014).
32. 'principles of protection must apply to any information concerning an identified or identifiable person: to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person' Recital (26).

data.[33] However, the Article 29 Working Party (Work Programme 2007) has considered pseudonymization and defined it as:

> ... the process of disguising identities. The aim of such a process is to be able to collect additional data relating to the same individual without having to know his identity. This is particularly relevant in the context of research and statistics.

Pseudonymization means that the individual is indirectly identifiable so data rules will apply but in the opinion of the Article 29 Working Party, 'the application of these rules will justifiably be more flexible than if information on directly identifiable individuals were processed'.[34] Pseudonymization is often used in research as it may be necessary to link the data with the data subject, for example, in a longitudinal cohort where different sources of data relating to an individual need to be linked over time.

In each of the BioSHaRE-EU jurisdictions, there are considerable differences in how these provisions are interpreted and what it means to pseudonymize data. This has significant implications for the transfer of data across borders as being able to use pseudonymized data is crucial for research conducted by large international consortia that operate across jurisdictional borders. The law provides limited guidance on what is required of researchers when they pseudonymize data.

The Finnish DPA does not contain specific rules regarding anonymized, coded or pseudonymized data. However, as the concept of personal data includes identifiable information relating to the family or household of an individual, this implies a higher level of safeguards will be necessary to reduce the risk of re-identification. The interpretation by the Data Ombudsman of Recital 26 of the Directive has been that any data that could be linked to a person is regarded as personal data, even if, for example, the identifier is administered by a third party. This sets a very high standard for researchers, even those taking reasonable precautions and leaving a low likelihood that data could be linked to the individual will be dealing with personal data.

In contrast to the Directive and the Finnish legislation that do not contain specific provisions on anonymization, pseudonymization or coded data, the German DPA defines the concept of 'rendering anonymous'. This is 'the alteration of personal data so that

33. It only recognizes that (within the meaning of Article 27) 'codes of conduct may be a useful instrument for providing guidance as to the ways in which data may be rendered anonymous and retained in a form in which identification of the data subject is no longer possible', see Recital (26) of the Data Protection Directive.

34. 'Indeed, using a pseudonym means that it is possible to backtrack to the individual, so that the individual's identity can be discovered, but then only under predefined circumstances. In that case, although data protection rules apply, the risks at stake for the individuals with regard to the processing of such indirectly identifiable information will most often be low, so that the application of these rules will justifiably be more flexible than if information on directly identifiable individuals were processed' (Article 29 Working Party, Opinion 4/2007 on the concept of personal data, 20 June 2007, WP 136 2007). Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf#page=15&zoom=auto,0,844 (accessed 2 February 2012).

information concerning personal or material circumstances cannot be attributed to an identified or identifiable natural person or that such attribution would require a disproportionate amount of time, expense and effort'.[35] However, this is another example of how general data protection rules may not be fitting for the protection of sensitive personal data in the context of biomedical research, for example, DNA is a unique identifier and therefore it is impossible to completely anonymize a sample.[36] It has been demonstrated that it is possible to identify individuals when genome-wide methodologies are used.[37]

Personal data collected or recorded for the purpose of scientific research may be processed or used only for these purposes and shall be rendered anonymous as soon as the research purpose allows. Until then, the features enabling the attribution of information concerning personal or material circumstances to an identified or identifiable person shall be kept separately. They may be combined with the information only to the extent required by the research purpose. Bodies conducting scientific research may publish personal data only if (a) the data subject has consented and (b) this is essential to present research findings concerning events of contemporary history. Personal data should not be transferred if the data subject has a legitimate interest in ruling out the possibility of transfer, especially if the bodies fail to ensure an adequate level of data protection.[38]

As in Finland, the UK DPA does not expressly refer to anonymized data, but it refers to the concept of identification or likelihood of identification through linkage of data in a similar way to the Directive. The Information Commissioner's Office Code of Practice on anonymization explains that 'this means that, although it may not be possible to determine with absolute certainty that no individual will ever be identified as a result of the disclosure of anonymised data, this does not mean that personal data has been disclosed'.[39] Recent case law has also addressed this issue in relation to statistical medical data.[40] In general, data no longer fall under the data protection principles, if through statistical disclosure control methods, the data are anonymized so that the chance of re-identification is very remote and the data can no longer be classed as personal.[41]

35. See section 3(6) of the German Data Protection Act.
36. J. Kaye, C. Heeney, N. Hawkins and J. de Vries, P. Boddington, 'Data Sharing in Genomics – Re-shaping Scientific Practice', *Nature Reviews, Genetics* 10 (2009), 334.
37. N. Homer, et al. 'Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-density SNP Genotyping Microarrays', *PLoS Genetetics* 4 (2008), p. e1000167; J. Couzin, 'Genetic Privacy. Whole-genome Data not Anonymous, Challenging Assumptions', *Science* 321 (2008), p. 1278.
38. See section 4b (1). Nevertheless, this rule should not apply if transfer is necessary for a public body of the Federation to carry out its duties for compelling reasons of defence or to fulfil supranational or intergovernmental obligations in the field of crisis management or conflict prevention or for humanitarian measures.
39. Information Commissioner's Office, 'Anonymisation: managing data protection risk code of practice', para. 16.
40. For example, *R (on the application of the Department of Health) v. Information Commissioner* [2011] EWHC 1430 (Admin).
41. Information Commissioner's Office (2012). Available at: http://www.ico.org.uk/for_organisations/data_protection/topic_guides/~/media/documents/library/Data_Protection/

As in Finland and in the UK, anonymization is not covered by Dutch DPA, but it is covered by the Civil Code. In situations where consent 'cannot reasonably be requested', researchers may obtain access to medical information or have it transferred to them for research provided that suitable guarantees are taken as regards the patient's privacy.[42] Anonymization is required so that 'data are supplied in such a form that they cannot be traced back to individual as to ensure that traceability to a natural person is reasonably prevented'.[42] The transfer of such information research must be in the public interest, the research cannot be carried out without the medical information from the patient's record and the patient must not have explicitly objected to such use of their medical record.[43] This appears to be in line with the safeguards generally present in health research aimed at reducing the risk of re-identification, but it does not take into account both its feasibility with regard to genetic data and the need for the researcher in certain cases to trace back the information to the individual patient. In particular, such a system may limit secondary uses of the data where specific contextual factors are needed.[44] However, such considerations have been taken into account in the federa codes,[45] which allow tracing back in certain instances, and by Article 10 of the Dutch DPA, which states that even if personal data shall not be kept in a form allowing the identification of the data subject for any longer than necessary for the purposes of the processing, it recognizes that they may be kept for longer where this is for historical, statistical or scientific purposes and where the responsible party has made the necessary arrangements to ensure that the data concerned are used solely for these specific purposes. This rule implements the principle contained in the Directive to allow further scientific data use where member states have developed 'suitable safeguards' (see Recital 29) but specifies it with regard to

---

Practical_application/anonymisation_code.ashx) (accessed 8 April 2014). See also R. Massey, 'Anonymisation: Managing Data Protection Risk – the New UK Code', *Computer and Telecommunications Law Review* 19 (2013), p. 86.

42. Article 7:458 of the Dutch Civil Code:

> Without prejudice to [existing obligations of secrecy] information about the patient or access to [medical files] may, if requested, be supplied to another person for the purpose of statistics or scientific research in the field of public health without the patient's consent, if a) consent cannot reasonably be requested and guarantees are provided that the patient's privacy will not be inordinately infringed by the conduct of the research; b) consent cannot reasonably be requested given the nature and purpose of the research and the care provider has ensured that the data are supplied in such a form as to ensure that they cannot be traced back to individual natural persons.

43. See Article 7:458.
44. S. Lindsay and J. Goldring, 'Anonymizing Data for Secondary Use', *Encyclopedia of Case Study Research* 2010, Available at: http://srmo.sagepub.com/view/encyc-of-case-study-research/n10.xml (accessed 2 february 2014); K.E. Emam, Methods for De-identification of Electronic Health Records for Genomic Research', *Genome Medicine* 3 (2011), 25.
45. Codes of Conduct for medical research set out by the Dutch Council of the Federation of Medical Scientific Societies (*Federatie Van Medisch Wetenschappelijke Verenigingen*). Available at: http://www.federa.org/gedragscodes-codes-conduct-en (accessed 8 April 2014).

data that may actually be traced back to the individual (see Article 10 para. 1).[46] As per a ruling of the Dutch Data Protection Agency, pseudonymization under certain conditions is deemed to render personal data anonymous and so has the effect of making the DPA no longer applicable to data thus pseudonymized. This ruling is heavily relied upon for purposes of health research.[47]

Unlike Finland, the Netherlands and the UK, which do not contain specific provisions regarding anonymization in their DPA, para. 20 of the Norwegian HRA states that consent is not required for research on anonymous human biological material and anonymous data, but it is required to collect material and data for subsequent anonymization.[48] As in the Dutch law, para. 38 further specifies that data must not be stored for longer periods than is necessary to complete the project, but the Norwegian Act also foresees the possibility of a different ruling by the regional ethic committee for medical and health research. The Committee:

> may rule that documents necessary for auditing the project must be kept for five years after the final report on the research project has been sent to the Committee, or for longer under certain conditions, and that if the data are not going to be kept thereafter in accordance with the Archives Act or other legislation, they must be anonymised or deleted.[49]

This rule, which refers in general to research using personal data, allows researchers the possibility of being able to have access to those data also for longer than necessary to complete the project, depending on a Research Ethics Committee (REC) decision. On the other hand, complete anonymization (when feasible) or deletion of the data may interfere with sharing data with researchers based in different jurisdictions, in particular for all secondary uses.

The latest draft of the proposed EU Data Protection Regulation contains a definition of 'pseudonymous data' under Article 4(2a), 'personal data that cannot be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution'. It is not clear whether pseudonymized data are intended to be included in the scope of the draft regulation, a concern raised by a coalition of European stakeholders who suggested that 'the Albrecht amendments would create a system in which the use of pseudonymised data is subject to most of the same regulatory

46. Article 7:458 of the Dutch Civil Code '... the care provider has ensured that the data are supplied in such a form as to ensure that they cannot be traced back to individual natural persons'.
47. Available at: http://www.medlaw.nl/wp-content/uploads/commentsdutchrshgdprv22.pdf at 2 (accessed 8 July 2014).
48. Cf. para. 20 of the Health Research Act under section 21. Research on human biological material from deceased persons. Research on biological material taken from deceased persons is correspondingly subject to the provisions in Act of 9 February 1973 no. 6 relating to transplantation, hospital autopsies and the donation of bodies and so on and regulations issued pursuant to this Act.
49. Para. 38, "Prohibition against storing unnecessary personal health data", Dutch Health Research Act, 2008-06-20 n. 44.

requirements as identifiable data'.[50] This could potentially increase the burden on relevant sectors of biomedical research, for example, large-scale population-based studies and cohorts involving biobanks and patient data.[51]

As pseudonymized data implies that the individual is still 'identifiable', it may be difficult to draw the line between what falls under the Data Protection Directive and what does not. The Data Protection Directive does not contain a definition of those terms, and our analysis shows that there are some differences across EU member states.

The lack of a unique definition has important implications for both donors and researchers since anonymized data do not fall under the scope of the data protection regulation and therefore may be freely accessible for research purposes. However, anonymization makes re-contact impossible. Such situation may raise issues for cross-border research, where the legal framework may vary from state to state, as well as in all the cases where re-contacting participants is fundamental to the progress of the research itself.[52,53]

Processing of personal data. Under Article 2(b) of the Directive 95/46/EC 'processing of personal data' means:

> 'any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction'.

---

50. Available at: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtvm054713.pdf (accessed 8 July 2014). Under Amendment 23 of the so-called Albrecht amendments:

> [. . .] This Regulation should not apply to anonymous data, meaning any data that cannot be related, directly or indirectly, alone or in combination with associated data, to a natural person or where establishing such a relation would require a disproportionate amount of time, expense, and effort, taking into account the state of the art in technology at the time of the processing and the possibilities for development during the period for which the data will be processed. (accessed 8 April 2014). Available at: http://www.janalbrecht.eu/fileadmin/material/Dokumente/DPR-Regulation-inofficial-consolidated-LIBE.pdf (accessed 2 Februaty 2015).

51. See 'Impact of the draft European Data Protection Regulation and proposed Amendments from the rapporteur of the LIBE committee on scientific research', March 2013, available at: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtvm054713.pdf (accessed 8 April 2014).
52. M. Barbareschi, S. Fasanella, C. Cantaloni and S. Giuliani, 'Scientific and Managerial Premises and Unresolved Issues in Tumor Biobanking Activities', in G. Pascuzzi, U. Izzo and M. Macilotti, eds. *Comparative issues in the governance of research biobanks.* (Berlin Heidelberg:Springer, 2013), pp. 301–309.; G. Laurie, P. Mallia, David A. Frenkel, A. Krajewska and H. Moniz, 'Managing Access to Biobanks: How Can We Reconcile Individual Privacy and Public Interest in Genetic Research?', *Medical Law International* 10(4) (2010), pp. 315–337.
53. J. Kaye, 'From Single Biobanks to International Networks: Developing E-governance', *Human Genet*ics 130 (2011), p. 377.

The Directive first sets out general conditions under which the processing of personal data is lawful (Article 7), leaving member states the possibility to determine these conditions more precisely. Then it prohibits the processing of special categories of data, like data concerning health, with the exception of certain cases. Data sharing is a form of data processing, as defined by the EU Directive 95/46/EC on data protection. Therefore, it is particularly important that data sharing in the context of biobanking research is based on uniform and harmonized requirements across member states.[54] Since funding bodies are increasingly requiring data sharing as a consideration of funding applications, this may be difficult to achieve if data processing rules as implemented by member states are slightly different one from each other.

Under Article 8(2), the processing of special categories of data, like data concerning health, will only be permitted if:

> (a) the data subject has given his explicit consent to the processing of those data, except where the laws of the Member State provide that the prohibition may not be lifted by the data subject's giving his consent; or (. . .) (c) processing is necessary to protect the vital interests of the data subject or of another person where the data subject is physically or legally incapable of giving his consent.

The processing will also be permitted if it:

> is required for the purposes of preventive medicine, medical diagnosis, the provision of care or treatment or the management of health-care services, and where those data are processed by a health professional subject under national law or rules established by national competent bodies to the obligation of professional secrecy or by another person also subject to an equivalent obligation of secrecy. (Article 8(3))

Subject to the provision of suitable safeguards, member states may, for reasons of substantial public interest, lay down exemptions in addition to those laid down in para. 2 either by national law or by decision of the supervisory authority (Article 8(4)). Section VIII of the Data Protection Directive also sets forth rules regarding the confidentiality and security of processing. Provisions regarding processing vary across the BioSHaRE-EU jurisdictions in slightly different ways.

The Finnish DPA implements the Directive rules without major differences.[55] The general principle is that the processing of sensitive data is prohibited, but some derogations are possible.[56] It follows the Directive but sets more detailed safeguards, which are far more specific also compared to the other BioSHaRE-EU countries.

---

54. B.M. Knoppers, J.R. Harris, A.M. Tasse, I. Budin-Liøsne, J. Kaye, M. Deschênes and M. Zawati, 'Towards a Data Sharing Code of Conduct for International Genomic Research', *Genome Medicine* 3 (2011), p. 46. Available at: http://genomemedicine.com/content/3/7/46
55. See chapter 3 of the Finnish Data Protection Act, section 12, para. (1)–(12).
56. Under para. 2, 'Sensitive data shall be erased from the data file immediately when there no longer is a reason for its processing, as provided in paragraph (1). The reason and the need for processing shall be re-evaluated at five-year intervals at the longest, unless otherwise

This could potentially complicate the creation of a common data sharing policy across biobanks or cohorts located in different jurisdictions since it must be compliant with the more specific Finnish requirements.

Section 14 (Chapter 4) of the Finnish DPA allows the processing of personal data for special purposes, with regard to historical or scientific research if the following specific conditions are fulfilled: (i) The research cannot be carried out without data identifying the person and the consent of the data subject cannot be obtained owing to the quantity of the data, their age or another comparable reason; (ii) the use of the personal data file is based on an appropriate research plan and a person or a group of persons responsible for the research have been designated; (iii) the personal data file is used and data are disclosed only for the purposes of historical or scientific research and the procedure followed does not disclose data pertaining to a given individual to outsiders; and (iv) after the personal data are no longer required for the research or for the verification of the results achieved, the personal data file is destroyed or transferred to an archive, or the data in it are altered so that the data subjects can no longer be identified. However, the Finnish DPA further clarifies that these provisions do not apply if this procedure is manifestly unnecessary for the protection of the privacy of the data subject owing to the age or quality of the data in the personal data file.[57]

Complementary and derogating provisions are also found in other specific Finnish laws. For example, the Act on the Status and Rights of Patients (the Patient Rights Act, 785/1992) regulates the use of patient records for research, that is, access to such records for research purposes is subject to a license granted locally by a local health care unit. Where several records are used from different regions, the National Institute for Health and Welfare may give permission to access this data.[58] There is a similar legislative situation in the other BioSHaRE-EU jurisdictions with the exception of Norway, where the HRA supersedes the DPA as well the other complementary regulations.

With regard to the processing of data by public bodies, the requirements in Germany are the same as the ones set out in the Data Protection Directive, but with an important difference regarding the form required for expressing consent. The processing shall be lawful when the knowledge of such data is necessary for the controller to perform its tasks and where (i) the data subject has given his/her consent in accordance with section 4a(3), required for the purposes of preventive medicine, medical diagnosis, the provision of care or treatment or the

provided in an Act or stated in a permission of the Data Protection Board referred to in'. See also Finnish National Ethics Committee (2005), 'Human Stem Cells, Cloning and Research'; Finnish National Ethics Committee, 'Opinion on EU research funding on stem cells', 28 February 2006; Finnish National Ethics Committee, 'Statement on Human Dignity and Rights of the Foetus', 13 April 2010.

57. See section 14, para. 2 of the Finnish Data Protection Act.

58. What is provided in the Act on the Openness of Government Activities, Patient Rights Act section 13, para. 4: Personal Data Registers for Health Care (556/1989) and in the Personal Data Act shall apply to the supplying of information contained in patient documents for scientific research and compilation of statistics. Patient Rights Act imposes secrecy on patient records and they shall not be disclosed to a third party without patient's consent. However, see section 28 of the Official Act. See also Act 795/2010 and the Act on National Personal Data Registers for Health Care (556/1989).

management of health-care services; (ii) these data are processed by health professionals or other persons subject to the obligation of professional secrecy; (iii) the processing is necessary for the purposes of scientific research, where the scientific interest in carrying out the research project significantly outweighs the data subject's interest in ruling out the possibility of collection and the purpose of the research cannot be achieved in any other way or would require a disproportionate effort. A unique feature of the German law compared to the other BioSHaRE-EU data protection laws, and which goes beyond the requirements of the Directive, is that consent for processing data shall be given in writing, unless special circumstances warrant any other form.[59] Consent shall be based on the data subject's free decision, who shall be informed of the purpose of collection, processing or use, as necessary in the individual case or on request, of the consequences of withholding consent.

The following special circumstances that may warrant forms other than writing occur: (a) in the field of scientific research, where the defined purpose of research would be seriously affected if consent were obtained in writing. Irrespective, it is necessary to record in writing for (i) the purpose of collection, processing or use, (ii) the consequences of withholding consent and (iii) the reasons the defined purpose of research would be seriously affected[60]; and (b) any other form other than written may be justified where special categories of personal data (section 3(9)) are collected, processed or used, that is, information on racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, health or sex life.[61] Despite these exceptions, requiring the written form as a general rule reflects the more restrictive approach to data protection rules found in Germany as compared to the other BioSHaRE-EU jurisdictions. Where consent is absent, the processing and use of personal data shall be lawful only if it is permitted or ordered by the Federal DPA or other law.[62]

Processing of personal data under the Dutch DPA reflects the rules contained in the EC Directive, as in the other BioSHaRE-EU jurisdictions. The general principle is that consent is required,[63] and it should be obtained respecting the relevant conditions. Under Article 16 of the Dutch Act, it is prohibited to process personal data concerning a person's health,[64] but this prohibition does not apply where the processing is (i) carried out with the express consent of the data subject, (ii) carried out without their express consent, but for the benefit of an important public interest, (iii) carried out where appropriate guarantees have been put in place to protect the individual's privacy and (iv) provided for by law, or the Data Protection Commission has granted an exemption. When granting an exemption, the Commission can impose rules and restrictions. It is possible to process personal data without expressing consent for the purpose of scientific research or statistics, if (i) the research serves a public interest, (ii) the processing is necessary for the research or statistics concerned, (iii) it appears to be impossible or would involve a disproportionate effort to ask for express consent and

59. If consent is to be given together with other written declarations, it shall be made distinguishable in its appearance. See section 4(1) of the German Data Protection Act.
60. See section 4a(2) of the German Data Protection Act.
61. See section 3(9) of the German Data Protection Act.
62. See section 4 of the German Data Protection Act.
63. See Article 23 of the Dutch Personal Data Protection Act.
64. See Article 21 of the Dutch Personal Data Protection Act.

(iv) sufficient guarantees are provided to ensure that the processing does not adversely affect the individual privacy of the data subject to a disproportionate extent.[65] In all these cases, the data may only be processed by persons subject to an *obligation of confidentiality* by virtue of office, profession or legal provision or under an agreement.

There are some provisions in the UK DPA that relate specifically to health/medical research and bypass the need for a data subject's consent. However, the general principle is that to have access to data it is necessary to obtain consent from the data subject and also to comply with the relevant conditions set out in the DPA, that is, (i) data are not processed to support measures or decisions with respect to particular individuals and (ii) they are not processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject.[66] The processing must comply with one of the conditions set out in Schedule 2 of the DPA. Under the Act, any data related to physical or mental health or condition is considered 'sensitive personal data'. For this kind of data, it is not necessary to obtain a data subject's consent (exemption from section 7),[67] if personal data are processed only for research purposes (which includes statistical or historical purposes) and (i) they are processed in compliance with the relevant conditions,[68] (ii) the results of the research or any resulting statistics are not made available in a form that identifies data subjects or any of them.[69] Any processing necessary for medical purposes, including medical research,[70] shall be undertaken by (a) a health professional

65. See Article 23.2 of the Dutch P Personal Data Protection Act. Moreover, under Article 44:

> 1. Where processing is carried out by institutions or services for the purposes of scientific research or statistics, and the necessary arrangements have been made to ensure that the personal data can only be used for statistical or scientific purposes, the responsible party shall not be required to provide the information referred to in Article 34 and may refuse to comply with the requests referred to in Article 35.

66. Section 33(3) of the UK Data Protection Act. Personal data that are processed only for research purposes in compliance with the relevant conditions may, notwithstanding the fifth data protection principle, be kept indefinitely.
67. Para. 7 of UK Data Protection Act, Right of access to personal data.
68. Schedule 2 and 3 UK Data Protection Act and para. 9 of the Data Protection Order 2000. The circumstances under which sensitive personal data may be processed are set out in the Data Protection (Processing of Sensitive Personal Data) Order 2000. In particular, see para. 9(b).
69. Section 33 of the UK Data Protection Act:

> (5) For the purposes of subsections (2) to (4) personal data are not to be treated as processed otherwise than for research purposes merely because the data are disclosed – (a) to any person, for research purposes only, (b)to the data subject or a person acting on his behalf, (c)at the request, or with the consent, of the data subject or a person acting on his behalf, or (d) in circumstances in which the person making the disclosure has reasonable grounds for believing that the disclosure falls within paragraph (a), (b) or (c).

70. Section 4(3) UK Data Protection Act, '(2) In this paragraph "medical purposes" includes the purposes of preventative medicine, medical diagnosis, medical research, the provision of care and treatment and the management of healthcare services'.

or (b) a person who in the circumstances owes a duty of confidentiality, which is equivalent to that which would arise if that person were a health professional.[71]

To meet the conditions for research, sensitive personal data may be processed where the processing is in the substantial public interest and does not support measures or decisions with respect to any particular data subject otherwise than with their explicit consent. Processing shall not cause, nor shall be likely to cause, substantial damage or substantial distress to the data subject or any other person. Personal data shall not be transferred to a country or territory outside the EU unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.[72] If data are anonymized, it can be transferred across borders without explicit consent.

As in the UK, there are also specific rules tailored to processing of health data in Norwegian law. Besides the main rule on consent expressed by para. 13 of the Act, chapter 7 specifically considers research using personal health data. The general principle is that consent must be obtained from participants in medical and health research, unless the law provides otherwise. It must be (i) informed, (ii) voluntary, (iii) express and documented and (iv) based on specific information about a concrete research project, unless there is a case for granting broad consent. Research participants may consent to human biological material and personal health data being used for 'specific, broadly defined research purposes'. This means that a broad consent model is expressly allowed under this new regime (see para 14). Processing of personal health data in medical and health research must comply with the conditions set forth in para 32 of the HRA:

i. it must have expressly indicated objectives,
ii. the health data must be relevant and necessary to achieve the objective of the research project,
iii. the degree of personal identification in the health data must not be greater than is necessary to serve the intended purposes and
iv. it may not be used for purposes that are incompatible with the original objective without the consent of the research participant, unless otherwise prescribed by law.

Researchers must obtain a new consent where the research project has changed substantially, if these changes are deemed to have consequences for the participant's consent. Otherwise the regional committee for medical and health research ethics may approve new or changed use of previously collected human biological material or

71. Schedule 3, para. 8 of the UK Data Protection Act.
72. Information Commissioner's Office, 8th Data Protection Principle, available at: http://ico.org.uk/for_organisations/data_protection/the_guide/the_principles (accessed 8 July 2014).

Such a rule does not apply if there is i) DS consent, ii) contract between DS and DC, iii) contract between data controller and third parties (related to DS/DC relationship), iv) Public interest, v) Legal proceedings, vi) Vital interest of DS, vii) Data is on a public record, viii) Uses EC – approved model contract, ix) Transfer is authorized by the EC. See schedule 4(1)–(9).

personal health data without new consent being obtained if (i) it is difficult to obtain it, (ii) the research in question is of significant interest to society and (iii) the participants' welfare and integrity are ensured.

In the latest version of the proposed EU Data Protection Regulation, draft Article 9 sets out a general prohibition of processing special categories of personal data and exceptions to this general rule, building on Article 8 of the Directive by adding genetic data. Under the current system, consent is one of the alternatives among equally valid bases for lawful processing (see Article 8.2[a]–[e] of the Directive), but under the proposed Regulation, processing of health data would be possible only with the consent of the data subject, under the conditions and safeguards referred to in Articles 81 and 83.[73] Members states would have to set an exemption of 'high public interest' for research without consent (Article 81(2)(a)). The new approach of the proposed EU Data Protection Regulation has raised a number of concerns, particularly, regarding its potential negative impact on biomedical and biobanking research. In fact, it leaves member states the possibility to adopt exceptions for research purposes, without harmonizing the area of scientific research where sensitive data are being used.

Under the proposed Regulation, processing of health data and therefore sharing data across biobanks will not only require the relevant consent from the data subject but also that certain conditions are fulfilled, including, as the proposed Article 83 requires, that 'data enabling the attribution of information to an identified or identifiable data subject is [sic] kept separately' from other information 'under the highest technical standards' and ensuring 'all necessary measures are taken to prevent unwarranted re-identification'. The amendment means that the exemption could only apply to the use of pseudonymized (or coded) data, even if research serves a high public interest. Moreover, the wording is not clear and leaves open a number of different interpretations of anonymized or pseudonymized data, and such legal uncertainty may hinder data sharing across different biobanks.[74]

Conclusions

This article identifies and compares the key legal requirements regarding research use of data in Finland, Germany, the Netherlands, Norway and the UK. In all the BioSHaRE-EU jurisdictions, in addition to the legislation implementing the EU Data Protection Directive, there are provisions for data protection in other regulations or sources of law. For example, in patient's rights legislation (Finland), hospital laws (Germany), civil codes (Germany and Netherlands), specific HRAs (Norway) or case law (UK). This fragmentation of the sources of law undermines the certainty of the rules necessary to guarantee the flow of the information contained in samples between biobanks and to allow researchers to share them without undue burdens. The proposed

73. Article 9.2 (h) and (i) of the proposed European Union draft Regulation.
74. M. Taylor and B. Thompson, 'Update on the European Data Protection Regulation', Available at: http://www.p3g.org/sites/default/files/site/default/files/Taylor%20Thompson%20-%20Data%20Protection%20Regulation%20Update%2025%20February%202014%20V3.pdf (accessed 2 February 2015).

draft of the EU Regulation on data protection will ensure a single set of rules on data protection across Europe, explicitly incorporating genetic information, but the latest amendments contain several changes that overall could have a negative impact on medical research.

The main grey areas of the 1995 Directive concern the concept of personal data, anonymization and data processing. The definition of personal data adopted in the BioSHaRE-EU jurisdictions leaves space for broad interpretation to also include genetic information as sensitive health data. In Finland, the definition of personal data of an individual expressly includes 'members of his/her family or household'. Arguably, this is in line with the views of the Article 29 Working Group and latest proposals for EU data protection reform. Anonymization and pseudonymization are not expressly regulated by the current Directive, which recognizes the possibility of member states adopting Codes of Practices on the issue. Even though the latest draft of the proposed EU Data Protection Regulation contains a definition of pseudonymization, it still leaves space for different interpretations of its scope.

With regard to data processing, the BioSHaRE-EU jurisdictions have implemented the Directive rules without major differences. The general principle is that the processing of sensitive data is prohibited unless express consent is provided or when it is for the purposes of historical, scientific or statistical research. A unique feature of the German law compared to the other BioSHaRE-EU countries, and which exceeds the requirements of Data Protection Directive, is that consent for processing data shall be given in writing.

As far as access to data is concerned, the general principle and the main relevant exceptions are the same in all BioSHaRE-EU jurisdictions; again, consent of the data subject is essential, but this requirement can be waived if access is requested for research purposes and if specific conditions are satisfied. However, these conditions vary from one country to another. The proposed EU Data Protection Regulation would require consent for processing health data for research purposes, leaving it to member states to introduce exceptions. If member states decide to introduce a research exemption, it could again lead to different legal requirements across the EU. This would affect international research consortia that need to share data and, to an extent, diminish the benefits of a single EU Regulation on data protection for biomedical research.

Our analysis has shown that although data protection rules across BioSHaRE-EU jurisdictions are substantially uniform and share the same European origin, there are still some grey areas relating to research that have not been clarified or specified at the national level. This leaves room for different interpretations by national courts. The situation is further complicated by the variety of regulations and sources of law through which EU data protection rules have been implemented in each member state, raising further problems linked to the hierarchy of the law. Importantly, even if a single EU Data Protection Regulation were enacted and directly applicable in all member states, as the current proposals stand, it would not provide enough clarification to remove the existing grey areas for data-based research. Unfortunately, the EU data protection law reform process seems to have taken insufficient account of the importance of data sharing for biomedical research and, ultimately, population health.

The peculiarities of biomedical research call for a specific regulation at European level, which could apply both for samples and data. In fact, there is a straight linkage between samples and data, that is, researchers require access to a sample in order to get the information contained in it, and the usefulness of biological samples for research is greatly reduced when associated data are not shared with the samples. Nevertheless, they are considered as two different legal entities as different sources of law regulate access to data and samples. Following the Norwegian example, having a unique regulation, applying both to samples and data, and specific, focusing on biomedical research, would clear the path towards simplification of the sources of law, guaranteeing a better harmonization across member states and the desired sharing of data across different jurisdictions.

The BioSHaRE-EU project tries to overcome the legal restraints by encouraging the development of cutting-edge technologies that enhance collaborations among investigators and enable the development of tools for data harmonization, database integration and federated data analysis.[75] In particular, it has proposed a new approach, named DataSHIELD,[76] a federated infrastructure allowing researchers to jointly analyse harmonized data whilst retaining individual-level data within their respective host institutions.[77] Such a system enables to share data for research purposes without the need to physically pool them together, therefore avoiding to incur in the traditional legal constraints.

The BioSHaRE-EU Healthy Obese Project aims to gain insights into the consequences of (healthy) obesity using data on risk factors and phenotypes across several large-scale cohort studies.[78] This project has combined 10 different cohorts in seven countries, using data transformed into a harmonized format, that is, it has used a computing infrastructure to enable the effective pooling of data and research into critical sub-components of the phenotypes associated with common complex diseases.[79]

---

75. D. Doiron et al., 'Data Harmonization and Federated Analysis of Population-based Studies: The BioSHaRE Project', *Emerging Themes in Epidemiology* 10 (2013), p. 12. Available at: http://www.webcitation.org/query.php?url=http://www.bioshare.eu&refdoi=10.1186/1472-6823-14-9 (accessed 2 February 2015).
76. (Data aggregation through anonymous summary statistics from harmonized individual level databases), M. Wolfson et al., 'DataSHIELD: Resolving a Conflict in Contemporary Bioscience – Performing a Pooled Analysis of Individual-level Data Without Sharing the Data', *International Journal of Epidemiology* 39(5) (2010), pp. 1372–1382.
77. D. Doiron et al., 'Data Harmonization'.
78. J.V.van Vlier-Ostaptchouk et al., 'The Prevalence of Metabolic Syndrome and Metabolically Healthy Obesity in Europe: A Collaborative Analysis of Ten Large Cohort Studies', *BMC Endocrine Disorders* 14 (2014), p. 9. Available at: http://www.biomedcentral.com/1472-6823/14/9 (accessed 2 February 2015).
79. Vlier-Ostaptchouk et al., 'Metabolic Syndrome'. The data harmonization and database federation methodology and infrastructure developed and piloted under BioSHaRE's HOP is founded on the DataSchema and Harmonization Platform for Epidemiological Research harmonization approach and on information technology tools developed by OBiBa (Open Source Software for BioBanks), see D. Doiron et al., 'Data Harmonization'.

BioSHaRE-EU has also supported the development of the so-called 'omics connect toolbox', which includes a set of software to be locally used by researchers for omics data validation, management, viewing and sharing.[80]

It has been created also a unified system to integrate and compare observation data across experimental projects, disease databases and clinical biobanks; the system comprises models, formats documentation and software that are all available for free and open source at http://www.observ-om.org.[81]

At the same time, BioSHaRE-EU cohorts do not have a common access policy, that is, researchers still need to contact each cohort individually to get access authorization. To overcome some of the issues raised by not having such 'common data access policy', an internal process has been developed, allowing tracking which cohorts under the BioSHaRE-EU project have provided access authorization for research.[82]

The BioSHaRE-EU project shows that cutting-edge technologies could enable effective sharing of data overcoming legal constraints and at the same time raises the question on whether different kind of governance systems are needed to support the current lack of a uniform regulatory framework.[83]

---

80. A. J. Brookes et al., 'Omics Connect Toolbox: A Management, Validation and Visualisation Toolkit for Complex Omics Data, European Society of Human Genetics conference 2013', poster, Available at: https://www.bioshare.eu/content/omics-connect-toolbox-management-validation-and-visualisation-toolkit-complex-omics-data (accessed 2 February 2015).
81. T. Adamusiak et al., 'Observ-OM and Observ-TAB: Universal Syntax Solutions for the Integration, Search, and Exchange of Phenotype and Genotype Information', *Human Mutation* 33(5) (2012), pp. 867–873.
82. Process and Data Access Schema developed by A. M. Tasse et. al. within the BioSHaRE-EU project. Data Access Procedure:

    > The Data Access procedure developed in BioSHaRE ensures that the research is feasible, does not conflict with other BioSHaRE research projects, and central governance of access to the data (Figure). 1.Investigators submit a research protocol to the Access Coordination Committee and the harmonization team, describing the research question, required variables and planned statistical analyses. The protocol is reviewed to assess potential overlap with ongoing research in the Core Project, to confirm the feasibility to harmonize the variables, and to perform analyses using DataSHIELD.

    2. Investigators request approval from their respective research ethics or data access committees to obtain the data. Approvals are registered and overseen centrally. 3. The Access Coordination Committee informs the harmonization team about the approved protocols and oversees investigator access to the data. 4. The harmonization team informs the central IT office and develops the processing algorithms. 5. The central IT office guides and instructs the local IT staff in installing the processing algorithms, providing access to registered investigators, updating software, etc. 6. The central IT office grants DataSHIELD access to registered investigators to use DataSHIELD on variables defined in their research protocol and only for the biobanks that approved the protocol." The approved protocols are published at the BioSHaRE website.
83. J. Kaye, 'Single Biobanks', pp. 377-382; J. Kaye, 'Do We Need a Uniform Regulatory System for Biobanks Across Europe?', *European Journal of Human Genetics* 14(2) (2006), pp. 245-248.

Acknowledgements

The authors are grateful for research done by Mr Nico Reiners into German data protection law and to Dr Jasper Bovenberg for having reviewed the Dutch legislation.

Funding

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 261433 (BioSHaRE-EU).