Biobank and Biomedical Research: Responsibilities of Controllers and Processors Under the EU General Data Protection Regulation

Ana Nordberg

Abstract Biobanks are essential infrastructures in current health and biomedical research. Advanced scientific research increasingly relies on processing and correlating large amounts of genetic, clinical and behavioural data. These data are particularly sensitive in nature and the risk of privacy invasion and misuse is high. The EU General Data Protection Regulation (GDPR) developed and increased harmonisation, resulting in a framework in which the specific duties and obligations of entities processing personal data—controllers and processors—were defined. Biobanks, in the exercise of their functions, assume the role of controllers and/or processors and as such need to comply with a number of complex rules. This chapter analyses these rules in the light of Article 89 GDPR, which creates safeguards and derogations relating to 'processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes'. It identifies key compliance challenges faced by biobanks as data controllers and processors, such as determining whether the GDPR is applicable and its intersection with other regulations; when a biobank should be considered controller and processor; and what are the main duties of biobanks as data controllers and processors and options for compliance.

1 Introduction

Biobanks, broadly understood, play a central role in contemporaneous medical and biomedical research. For its part, scientific biomedical research is essential in modern developed societies and serves the realisation of important fundamental rights, namely the right to life and health care.[1] Cutting-edge health research increasingly

[1] Article 2 'Right to Life'; Article 35 'Right to health care', Charter of Fundamental Rights of the European Union *OJ C 326, 26.10.2012, p. 391–407.*

A. Nordberg (✉)
Lund University, Faculty of Law, Lund, Sweden
e-mail: ana.nordberg@jur.lu.se

S. Slokenberga et al. (eds.), *GDPR and Biobanking*, Law, Governance and Technology Series 43, https://doi.org/10.1007/978-3-030-49388-2_5

relies on large amounts of genetic, clinical and behavioural data. These data are particularly sensitive and enjoy increased legal protection,[2] thus creating complex intersections between fundamental values. Data protection law has a long history in Europe, and unlike other jurisdictions such as the USA it is based on the principle that personal data processing is prohibited unless explicitly allowed under a specific legal basis.[3] The latest data protection development in the EU is the GDPR,[4] which replaced the previous framework set forth by the Data Protection Directive.[5]

The present chapter focuses specifically on the duties of biobanks as data controllers and data processors under the GDPR. The GDPR has created an increasingly harmonised framework as to the duties and obligations of entities which retrieve, store and analyse personal data, i.e. data controllers and data processors. Biobanks, in their typical operating functions, assume the roles of controllers and processors of personal data. From the perspective of biobank compliance with the duties and obligations imposed by EU data protection law, relevant key changes include: (1) higher penalties for contravention; (2) new requirements for appointment of a data protection officer (DPO) when an entity processes significant amounts of sensitive data; (3) recognition of genetic data as sensitive personal data; (4) strong promotion of a privacy by design approach; (5) new direct obligations imposed on data processors; (6) broader territorial scope, now expanding to non-EU entities which process EU citizens' data; (7) time limitation on the storage of data; (8) specific permission for broad consent for scientific research; (9) exemption from some individual data subject rights concerning data 'for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes'.[6]

Whether or not biobanks assume the roles of data controllers and/or data processors for GDPR compliance purposes will largely depend on their actual functions, manner of operating and whether the specific tasks can be considered data processing of personal data. In order to contextualise the debate on the duties of biobanks as data controllers and processors, it should be briefly mentioned that data protection rules intersect with the general regulatory frameworks applicable to biobanking activities in the EU and EU Member States. Among the European biomedical community, biobanking terminology tends to vary.[7] There is therefore

[2] The right to privacy is a fundamental right linked to the notions of human dignity, equality and autonomy. See for example Article 7 'Respect for private and family life; Article 8 'protection of personal Data'; Article 21 'Non-discrimination' EU Charter of Fundamental rights.

[3] Dove (2019).

[4] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, *OJ L 119, 4.5.2016, p. 1–88.*

[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, 23.11.1995 P. 31 – 50.*

[6] Article 9(2)(j) GDPR; Morrison et al. (2017), pp. 693–703.

[7] Fransson et al. (2015), pp. 22–28; Watson (2014), pp. 163–164; Hewitt and Watson (2013), pp. 309–315; Shaw et al. (2014), pp. 223–227.

neither a common understanding of what a biobank is nor agreement on a taxonomy of different types of biobanks. Legislation across EU Member States reflects the difficulties in establishing precise legal definitions of biobanks and biobanking activities.[8] At the national level, regulative approaches to biobanks reflect the pluralism of ethical, research and legal traditions and have their roots in significant socio-political, cultural and religious normative diversity.[9] Only a minority of EU Member States have specific legislation on biobanks.[10] The majority either do not have any domestic legislation[11] or rely on non-specific existing laws, often accompanied by soft law instruments, such as ethical guidelines, to regulate biobanks.[12] Lack of EU harmonisation and diversity of solutions, and in some cases vague and dispersed legislation, are all considered problematic for the development of biobanking activities.[13]

Overall, biobanks are quite diverse in terms of features such as the number, type and nature of samples, population covered, type of associated information, purpose and activities developed (e.g. sample hosting, processing and curation). These specific features influence the intersections between legal regulation of biobanking activities (mainly national) and the EU data protection framework and have practical implications for compliance with the obligations imposed by the GDPR on controllers and processors of personal data. There is a lack of specific, harmonised EU legislation on biobanks and biobanking activities. Existing EU regulation applicable to biobanks and biobank research is dispersed through a number of areas of law, including data protection, clinical trials[14] and tissue regulation.[15] An exhaustive analysis is outside the scope of this chapter. However, it can be noted for example the complex interplay between clinical trials regulation and the GDPR.[16]

---

[8] Beier and Lenk (2015), pp. 69–81; Briceño Moraia et al. (2014), pp. 187–212.

[9] Penasa et al. (2018), pp. 241–255.

[10] Belgium, Estonia, Finland, Hungary, Latvia, Lithuania, Portugal, Spain, Sweden and UK.

[11] Bulgaria, Croatia, Czech Republic, Malta, Romania, Slovakia.

[12] Austria, Cyprus, Denmark, France, Germany, Greece, Italy, Luxembourg, the Netherlands, Poland, Republic of Ireland and Slovenia. See Beier and Lenk (2015). See also: Nicola (2015), pp. 800–815; Sandor et al. (2009).

[13] In this sense, see a for example Penasa et al. (2018), with further references to national commentators defending the introduction of specific codified legislation in their respective jurisdictions.

[14] Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use, OJ L 121, 1.5.2001, p. 34, soon to be replaced by entry into effect of Regulation 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC, OJ L 158 27.05.2004, p. 1–76 [hereinafter Clinical Trials Regulation].

[15] Directive 2004/23/EC of the European Parliament and of the Council of 31 March 2004 on setting standards of quality and safety for the donation, procurement, testing, processing, preservation, storage and distribution of human tissues and cells, *OJ L 102, 7.4.2004, p. 48–58.*

[16] See European Data Protection Board, Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR) (art. 70.1.b)), Adopted on 23 January 2019.

This chapter examines the obligations imposed by the GDPR on biobanks in their role as controllers or processors of human personal data. After this introduction which sets out the contextual background of the application of data protection norms to biobanking activities, Sect. 2 addresses the material and geographic scope of applicability of the GDPR concerning biobanking activities. Section 3 then examines the concepts of controller and processor, their relationships and how these apply in a biobanking context. Section 4 analyses the duties of biobanks as data controllers and processors by reference to general data processing principles and the related duties imposed on biobanks, including obligations to respect data protection rights of data subjects. Adopting the perspective of biobanks as controllers and processors of data, it addresses possible compliance routes, with particular emphasis on rules concerning data processing of health and genetic data and exemptions provided for data processing 'for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes'.[17] Section 5 will conclude this chapter with a general summary of the main points addressed.

2 GDPR and Biobanking Activities

2.1 Substantive Scope of the GDPR

Data protection obligations of biobanks depend largely on their geographical establishment, location of data subjects, functioning, tasks performed and whether these allow their classification as controllers and/or processors of personal data under the EU jurisdiction. In other words, in order to determine whether in a specific situation a biobank has to comply with the GDPR rules, it is necessary to establish whether it falls both under the substantive and the geographic scope of application of the Regulation.

In substantive terms, the GDPR applies to data processing activities and these are defined broadly and generally, which means that in practice they will include most biobanking activities and related research. Any activity involving personal data, performed either by automated or manual means, is in principle subject to the GDPR. This includes, for example, 'collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction'.[18]

Data protection rules only apply to personal data, which means information relating to an identified or identifiable living, natural person. The concept of identifiable natural person is broadly defined and identification does not need to be immediate and direct. Data will still be personal if an individual can be identified by

[17] Article 9(2)(j) and Article 89(1) GDPR.

[18] Article 4 (2) GDPR.

reference to an identifier, for example, name, number, IP or physical address, or specific physical, physiological, genetic, mental, economic, cultural or social descriptors.[19] The concept of personal data only applies to living persons, and therefore *prima facie* it will not apply to samples obtained from deceased individuals. However, personal data of living relatives can be inferred from historical samples, thus arguably when inferences are established concerning, for example, the health of a living relative, such might constitute personal data processing under the GDPR.

2.2 Geographical Scope of the GDPR

Biobanks often collect, receive, keep or analyse transnational samples or data, which raises the question of the geographic scope of applicability of data protection rules. Generally, there are two factors that are relevant to determine the territorial scope of application: the establishment criterion, and the targeting criterion.[20] These will be further examined below.

Concerning the establishment criterion, the European Data Protection Board (EDPB) recommends consideration of three aspects: (a) establishment in the EU; (b) processing of personal data carried out 'in the context of the activities of' an establishment; and (c) application of the GDPR to the establishment of a controller or a processor in the EU regardless of whether the processing takes place in the EU or not.[21] The GDPR has a broad scope of applicability as it does so regardless of where the data processing activities are conducted and to any processing of personal data done by a controller or a processor with an establishment in the EU.[22] Recital 22 clarifies that 'establishment implies the effective and real exercise of activity through stable arrangements'.[23] Factual elements and not legal formalities are the determining factor to assess whether a data controller or processor has an establishment in the EU. In some circumstances, the GDPR rules also apply even if the controller or processor is not established in the EU as long as the data subject is located in the EU. In a biobank context, whether the data processing is considered carried out in the context of the activities of an establishment does not depend necessarily on whether the processing in question is carried out 'by' the biobank itself.[24] Assessment will have to be made on a case by case basis. For example, in cases of data and sample sharing, the activities of a biobank in a Member State and the data processing activities of a third party (data controller or processor) outside the EU may be inextricably linked, and thereby may

[19] Article 4 (1) GDPR.

[20] EDPB Guidelines 3/2018 on the territorial scope of the GDPR (Article 3) Adopted on 16 November 2018.

[21] EDPB Guidelines 3/2018, p. 4–7.

[22] Article 3 (1) GDPR.

[23] Recital 22 GDPR.

[24] Article 3(1)

trigger the applicability of EU data protection law even if the biobank by itself does not have an active role in the data processing.[25] Finally, the place of processing is not relevant in determining whether or not the data processing, carried out in the context of the activities of an EU biobank, falls within the scope of the GDPR. For example, when samples and information are collected outside the EU and later the data are processed by a biobank operating in an EU Member State or when a clinical trial is conducted outside the EU by a branch or subsidiary not legally distinct from an EU entity which determines the purpose and means of the data processing carried out on its behalf.[26]

In regards to the targeting criterion, Article 3 contains international private law rules that extend the jurisdiction of the GDPR to data controllers and processors not established in the EU and regardless of where the data processing activities take place. The connecting factor here is the location of the data subject and the purpose of the data processing activities. The GDPR applies to data subjects located in the EU[27] independently of their legal status concerning nationality or residence.[28] The second cumulative jurisdiction connecting factor concerns the type of data processing activities. Article 3(2) GDPR defines these as:

(a) the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or
(b) the monitoring of their behaviour as far as their behaviour takes place within the Union.

Biobanking activities may involve offering goods or services, such as where tissues and living materials are preserved as a service, for example, preservation of stem cells present in the umbilical cord or preservation of gametes and embryos for future use in an IVF context. The EDPB considers that it is necessary to have an actual 'connection between the processing activity and the offering of good or service, but both direct and indirect connections are relevant and to be taken into account'.[29]

The second type of activity that triggers the application of the GDPR to controllers or processors not established in the EU is the monitoring of data subject

[25] EDPB Guidelines 3/2018. See: Judgment of the Court (Third Chamber) 1 October 2015, Case C-230/14, Weltimmo s. r. o. *v* Nemzeti Adatvédelmi és Információszabadság Hatóság, Digital Reports: ECLI:EU:C:2015:639 para. 25, and Judgment of the Court (Grand Chamber), 13 May 2014, *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos* (AEPD) *and Mario Costeja González*, Case C-131/12, *Digital reports:* ECLI identifier: ECLI:EU:C:2014:317, para. 5.3.

[26] Adapted from EDPB Guidelines 3/2018, p. 8.

[27] Article 3 (2) GDPR, see also Article 8 EU Charter where the right to data protection is not limited to 'citizens but intended for 'everyone'.

[28] Recitals 2, 14 and 24 GDPR.

[29] EDPB Guidelines 3/2018, p. 21. see also Recital 23 GDPR and CJEU case law based on Regulation 44/2001 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters, for example, Joined Cases C-585/08 and C-144/09: Judgment of the Court (Grand Chamber) of 7 December 2010 (references for a preliminary ruling from the Oberster Gerichtshof (Austria))—Peter Pammer v Reederei Karl Schlüter GmbH & Co KG (C-585/08) and Hotel Alpenhof GesmbH v Oliver Heller (C-144/09), *OJ C 55, 19.2.2011, p. 4–5.*

behaviour as far as their behaviour takes place within the Union.[30] These are two cumulative criteria. The nature of the processing activity that can be considered as behavioural monitoring is further specified in Recital 24, which focuses exclusively on the monitoring of a behaviour through the tracking of a person on the internet. However, the EDPB considers that tracking through other types of network or technology involving personal data processing should also be taken into account, for example, through wearable and other smart devices. In a biobanking research context, monitoring may occur in longitudinal studies involving multiple samples and health information retrieved over time or where data subject information is regularly updated. However, it is not clear whether this represents behaviour monitoring since the spirit of the GDPR elucidated in Recital 24 GDPR clearly points to commercial monitoring of consumers. Regardless of this, since health and genetic data enjoys additional protection, there is good reason to understand that health monitoring can also be included and will thus trigger the application of the GDPR.

3 Notion of Controller and Processor in Biobanking

3.1 Definition of Controller and Processor

In the GDPR, the duties of data controllers and processors have been framed as positive obligations which emanate from the individual rights of data subjects,[31] for example, the rights to information, access, rectification, erasure and blocking, and to object to the processing of personal data. From a compliance perspective, this means that the first and foremost important task is to ensure a full understanding of the role each intervenient in biobanking research assumes for data protection purposes.

The legal concepts of controller and processor are established in Article 4 (7) and (8) GDPR as follows:

> 'controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law;
>
> 'processor' means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller;

These definitions have been transplanted without modification from the Data Protection Directive[32] and have their origin in a similar text in the Council of

[30] Article 3(2)(b) and Recital 24 GDPR.

[31] See Chapter, Staunton C (2019) Individual rights in Biobank research under the GDPR.

[32] *Directive 95/46/EC*. The concept of 'controller' was adopted with a few modifications from the Council of Europe's Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, Strasbourg, 28/01/1981 (CoE ETS 108).

Europe's Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data[33] concluded in 1981. Although the wording appears relatively straightforward, in practice it may not be so simple to assert who is the entity responsible for determining the purposes and means of data processing and identify the (various) entities processing data on behalf of a controller. This is due to contemporaneous organisational differentiation and complexity in both the public sector and private industrial fabric. The scope of these concepts was clarified by Opinion 1/2010 of the Article 29 Data Protection Working Party (WP29).[34] This soft law instrument analysed each operative concept of the definitions or its three main building blocks: (1) the personal aspect; (2) the possibility of pluralistic control; and (3) the essential elements to distinguish the controller from other actors—'determination' of 'purpose' and 'means'.[35] Controller and processor are independent functional EU concepts to be concretely determined by reference to the factual reality. This means that the type of activities of a biobank will have a bearing on whether and what entities are considered controllers and processors.

A controller is defined by its function and ability to decide on the purposes of processing and the means used. This role is based on a notion of control which can stem from any form of legal entitlement, including both explicit and implicit legal competence or from factual influence. The controller is also defined by its ability to determine the substantive content of the data processing. This ability must not be absolute: there is room for discretion and delegation. Whoever makes a de facto determination of the 'purpose' of processing is a controller while concrete methodologic issues concerning the choice of 'means' of processing can be delegated. In short, in a biobanking context the controller is whichever entity decides on issues pertaining to those substantial questions which are essential to the core of lawfulness of processing, for example, decisions on the legal basis for processing (e.g. consent or an exception), length of time a biological sample and related data are to be stored and who has access to the personal data processed.

The concept of processor is dependent on the organisational decisions and structure of the controller. The GDPR establishes two basic conditions for qualifying as processor: being a separate legal entity and processing data on behalf of a controller. Since, the controller decides either to process data within the organisation or to delegate all or part of the processing activities to an external entity, generally, processing data 'on behalf' means serving someone else's interest and is linked to the general concepts of 'delegation' and 'representation'. A processor implements

[33] CoE ETS No.108. This convention, was the only international legally binding instrument on the protection of private life and personal data open to any country in the world, and has been revised by the Protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (CoE ETS No. 223), 128th Session of the Committee of Ministers, Elsinore, 17–18 May 2018.

[34] Opinion 1/2010 issued by reference to the data protection directive, remains valid since these definitions transited unchanged to the GDPR.

[35] Opinion 1/2010.

instructions and decisions of the controller at least with regard to the purpose of the processing and the essential elements of the means.

3.2 Joint-Controllers and Joint-Processors

Data processing responsibilities may be borne by any natural or legal person and if shared will give rise to the notion of joint-controllers and joint-processors. In biobanking practice, situations involving putative joint-controllers and joint-processors present challenges, in particular when different entities submit samples and data to a biobank and/or when such data are shared, used and re-used by a diverse number of research institutions. The jurisprudence of the CJEU supports a broad concept of controller. In *Wirtschaftsakademie*[36] the Court of Justice of the EU (CJEU) ruled on joint-controllers, reaffirming the broad concept of controller previously established in *Google Spain.*[37] The court based its ruling on the criteria of whether a processor contributes, in the specific context, to determining, jointly with the main controller, the purpose and means of processing the personal data.[38] Applying this reasoning to a biobanking research context, both biobanks, researchers and entities conducting, sponsoring or financially supporting research, may be considered data controllers either by themselves or jointly. Their role differentiation and attribution will depend on the contractual relationships and de facto organisation of the research activities. Any entity which processes data on behalf of the controller will be considered a data processor. These activities comprise 'collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction'.[39]

3.3 Relationship Between Controllers and Processors

Controllers are responsible to ensure that those entities that process the data comply with data protection rules. Contractual relationships established between biobanks and research institutions or commercial companies should set up an allocation of tasks, rights and obligations between the parties, including provisions concerning

[36] Judgment of the Court (Grand Chamber) of 5 June 2018, *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH*, Case C-210/16, OJ C 260, 18.7.2016, ECLI:EUC:2018:388.

[37] Judgment of the Court (Grand Chamber), 13 May 2014, *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos* (AEPD) and Mario Costeja González, Case C-131/12, *Digital reports:* ECLI identifier: ECLI:EU:C:2014:317.

[38] C-210/16 *Wirtschaftsakademie*.

[39] Article 4(2) GDPR.

the purpose of processing, type of personal data and categories of data subject involved. Among other specific subjects, data processing contracts should address the issue of transfers of data to countries outside the EU or to international organisations.[40] Contracts should also include clauses on subcontracting of data processing activities as processors are precluded from subcontracting without the controller's prior written agreement.[41]

Territorial scope is also relevant here as often biobanking activities are conducted in collaboration with international research institutions and repositories. Firstly, the EDPB takes the view that the existence of a relationship between a controller and a processor does not necessarily trigger the application of the GDPR to both if one is not established in the Union. This means that 'when it comes to the identification of the different obligations triggered by the applicability of the GDPR, the processing by each entity must be considered separately'.[42] Secondly, when an EU biobank acting as a controller uses a processor located outside the EU, it will be necessary for the controller to ensure by contract or other legal act[43] that the processor will conduct its activities in accordance with the GDPR. This will include imposing on the processors by contract clauses all the relevant obligations placed by the GDPR on processors, and thus extending by contractual means the GDPR scope of application to processors outside the EU. Thirdly, the opposite situation—a biobank processing data on behalf of an institution/controller outside of the EU—is also a recurrent one. In such cases, while the provisions of the GDPR do not apply to the data controller, the biobank, as a processor established in the EU, will still continue to be required to comply with the GDPR obligations imposed on data processors provided that such activities are carried out in the context of its activities.[44]

4 Duties of Biobanks as Controllers and Processors

4.1 Accountability

Biobanks are responsible and accountable for compliance with data protection rules in their various activities as data controllers, for example, in receiving, holding or distributing biological samples or materials and associated data.[45] This means that biobanks in their capacity as data controllers are responsible for implementing the appropriate technical and organisational measures both to ensure compliance and to be able to demonstrate compliance with GDPR principles and rules.

[40] Article 26(3) GDPR.

[41] Article 26(2) GDPR.

[42] EDPB 3/2018, p. 9.

[43] Article 28(3) GDPR.

[44] EDPB 3/2018, pp. 10–11.

[45] Article 5(2) GDPR.

As seen above, the accountability obligations of biobanks also include exercising a supervisor function and ensuring that researchers and entities in the position of personal data processors follow data protection rules.[46] If several entities are in the position of data controller, they become joint-controllers. For reasons of legal certainty, joint-controllers have the additional responsibility to determine in a transparent manner the allocation of the shared responsibilities for compliance.

Data protection rules establish the rights of data subjects and impose corresponding duties on controllers and processors. These comprise both the general duty to assure compliance with general principles of data protection stemming from the principle of accountability and specific duties pertaining to the factual relationship and conduct towards data subjects in the course of data processing activities. General data protection principles include: (1) lawfulness, fairness and transparency; (2) purpose limitation; (3) data minimisation; (4) accuracy; (5) storage limitation; and (6) integrity and confidentiality.[47]

The principle of 'accountability' inverts the burden of proof, imposing on biobanks acting in the capacity of data controllers the responsibility for demonstrating that all data processing activities are conducted lawfully, fairly and in a transparent manner in relation to the data subject.[48] 'Lawfulness' of data processing activities is the fundamental basis for compliance with all other duties of controllers and processors under EU data protection law. If data are processed unlawfully, compliance with other duties and obligations will not preclude eventual sanctions. This means that, in the absence of legitimate grounds for data processing, all ensuing biobanking activities will be tainted by the unlawfullness of data processing. Because the right to data protection and privacy are fundamental rights protected by the EU Charter, the legal consequences of unlawful data processing may even expand beyond data protection sanctions. For example, it may hinder the ethical acceptance of the research for patentability purposes.[49] Once lawfulness of processing has been established, biobanks and biobank researchers will have to ensure effective compliance with the other principles of data protection mentioned above and the associated duties imposed on data controllers and processors. 'Purpose limitation' means that personal data can only be processed for specified, explicit and legitimate purposes. Further processing outside the initial purpose/conditions is generally not allowed. An exception is made for 'processing for public interest, scientific or historical research or statistical purposes'.[50] 'Data minimisation' means that processing activities are required to be adequate and relevant to the purposes, and the privacy intrusion is limited to the minimum necessary to achieve such purposes.[51] The principle of accuracy imposes the duty to take reasonable steps to ensure that inaccurate or

[46] Article 28(1) GDPR.

[47] Article 5 GDPR.

[48] Article 5(2) GDPR.

[49] Nordberg and Minssen (2016), pp. 138–177; Hellstadius and Schovsbo (2018).

[50] Article 5(1)(b) GDPR.

[51] Article 5(1)(c) GDPR.

out of date information is rectified or erased.[52] 'Storage limitation' refers to the duty to anonymise or erase data once it is no longer necessary for achieving the original purposes. This principle is also an object of limitation if the personal data are processed solely 'for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes' provided that the processing is subject to appropriate technical and organisational measures to safeguard the rights and freedoms of data subjects.[53] Finally, 'integrity and confidentiality' of personal data against unauthorised or unlawful processing, as well as accidental loss, destruction or damage, is to be ensured by the use of appropriate technical or organisational measures.[54]

4.2 Lawfulness of Data Processing

4.2.1 Categories of Personal Data and Lawfulness in Biobanking

It is critical to consider data types and their relevance for determining the concrete duties and compliance obligations of data controllers and processors. Unlike data subjects, not all personal data are born equal. Some types of informational content are liable to cause greater intrusion in the data subject's personal private sphere and/or have a higher risk of being misused for discriminatory practices or outcomes. The rapid development and availability of DNA sequencing, big data techniques and artificial intelligence (AI) has in recent years changed biomedical research and biobanks. Biological samples are now accompanied by personal data that can be aggregated and correlated through data mining techniques in a variety of ways. Such personal data may originate from health and medical records but also from research and clinical trials and other sources. It may include genetic and genomic data and other epistemological biomedical information but also environmental, lifestyle or social data.

As mentioned, processing personal data is only allowed under specific grounds and stricter rules apply concerning processing of special categories of personal data, including health data and genetic data.[55] It is therefore important, as a matter of compliance, that biobanks distinguish between non-personal and personal data but also between general personal data and special categories of personal data.

The concept of health data is defined in the GDPR as 'personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status'[56] and this includes 'all data pertaining to the health status of a data subject which reveal information

[52] Article 5(1)(d) GDPR.

[53] Article 5(1)(e) GDPR.

[54] Article 5(1)(f) GDPR.

[55] Article 9 GDPR.

[56] Article 4(15) GDPR.

relating to the past, current or future physical or mental health status of the data subject'.[57] Health data include both information derived from health records and 'information derived from the testing or examination of a body part or bodily substance, including from genetic data and biological samples'.[58]

Genetic data means 'personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or the health of that natural person and which result, in particular from an analysis of a biological sample from the natural person in question',[59] in particular, 'chromosomal, deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) analysis, or from the analysis of another element enabling equivalent information to be obtained'.[60] The GDPR imposes obligations on data controllers and processors with a focus on regulating data processing from the perspective of lawfulness of such processing. However, it does not regulate what types of derivative information can be obtained (correlations and inferences) nor what types of uses of data are permissible. Particularly problematic data uses, such as predictions and correlations based on big data analytics and AI, are still only timidly regulated.[61] The type of research activities developed by each biobank will have a bearing on determining the most suitable legal basis to rely upon for compliance with the principles of lawfulness, fairness and transparency. In any case, this decision must be made beforehand since controllers have the duty to inform individual sample donors/owners of the legal grounds allowing the data processing before collecting or in any way processing data.[62] Because new data processing technologies such as big data analytics allow category jumping inferences, it will often be the case that all data will become personal data, if not immediately then at least in the future. Moreover, the use of biological samples will equate to actual or potential genetic data and health data, and thus a cautionary approach would lead to generally considering that most data processed by biobanks and biobanking research are likely to pertain to one of the special categories of personal data.

4.2.2 Modalities for Lawful Data Processing in Biobanking

General Remarks

Ensuring the lawfulness of data processing is the most essential duty of controllers and processors. In ensuring lawfulness, choosing an appropriate legal basis for processing the data is of utmost importance and has to be performed prior to the

[57] Recital 35 GDPR.

[58] Recital 35 GDPR.

[59] Article 4(13) GDPR.

[60] Recital 35 GDPR.

[61] Article 22 GDPR.

[62] Article 7 GDPR.

collection of data. The GDPR contains several legal basis for data processing. These can be conceptualized as two main models for lawfulness of data processing in biobanks and bio-banking research: (a) consent-based model, and (b) necessity-based model. Depending on the ground for lawfulness, different obligations will be imposed on biobanks in the capacity of either data controllers or data processors. In order to simplify the compliance analysis, in this section it will be assumed that most human data processed by biobanks or in biobanking research are special categories of personal data (e.g. health data and genetic data), and thus attention will focus on the lawfulness grounds established in Article 9 GDPR.

Necessity-Based Model

Generally, the processing of special categories of personal data, such as genetic and health data, is prohibited. However, biobanks can choose to rely on the exceptions and exemptions provided in Article 9(2) GDPR and so implement either a consent or necessity-based model or a mixture. Among the various exceptions conferring lawfulness of processing, of particular interest for biobanks is data processing justified by the necessity 'for archiving purposes in the public interest, scientific or historical research purposes'[63] and processing justified by the necessity 'for reasons of public interest in the area of public health'.[64] This data processing model can be suitable where obtaining consent is not possible or excessively burdensome (for example, when data is re-purposed and contact information is missing or outdated), or when consent is insufficient, redrawn or denied. The definition of 'scientific research purposes' is broadly constructed and includes 'technological development and demonstration, fundamental research, applied research and privately funded research'.[65]

In biobanking research, re-use and repurposing of data has become a necessity where new digital technologies offer increased possibilities to cross-reference large quantities and types of data from multiple sources (big data analytics), including health and medical records. However, data have to be collected 'for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes'.[66] This means that the lawfulness of data processing has to be established prior to the data collection and is connected to the purpose for which the data were collected. The result of this is that a necessity-based model may offer advantages to biobanks and in certain circumstances be the preferred option to establish lawfulness since repurposing of data for archiving or research purposes is generally presumed compatible with the original purpose as long as the controller

[63] Article 9(2)(j) GDPR.

[64] Article 9(2)(i).

[65] Recital 159 GDPR. The recital mentions specifically 'studies conducted in the public interest in the area of public health'.

[66] Article 5(1)(b) GDPR.

demonstrates respect for individual rights and freedoms of the data subject and implements appropriate safeguards, such as pseudonymisation (unless this is impossible or impairs the archiving or research purposes).[67]

Under Article 9(2)(j), processing of health and genetic data without consent is possible for scientific research purposes provided that processing is: (a) necessary for scientific research purposes; (b) proportionate to the aims pursued; (c) and respects the essence of the right to data protection.[68] These requirements will be relatively simple to fulfil in the case of biobanking activities directly connected with a specific research project aimed at studying a serious medical condition. However, concerning biobanking activities not directly linked to a specific research project or where such a link is less immediate or evident, data controllers will need to carefully justify that the use of the data is necessary and proportionate. In any case, the essence of the data protection right must be respected. This means that all processing activities must respect the general principles of data protection: lawfulness, fairness and transparency; purpose limitation; data minimisation; accuracy; storage limitation; integrity and confidentiality; and accountability.

Article 9(2)(i) GDPR allows Member States to establish the lawfulness of data processing for public interest reasons in the area of public health. Provided that a legal basis exists and specific measures to safeguard the rights and freedoms of data subjects and the confidentiality of health records are enacted, samples collected in the course of medical treatment might be stored in biobanks and made available for research, alongside patient records. However, a non-consenting data subject is unlikely to collaborate and provide additional samples or necessary specific information, thus affecting the ability to monitor an individual's health over time or study the health impact of specific lifestyle or social and environmental factors. Because patient records, even if standardised and comprehensive, are often of limited interest to researchers, the consent-based model will remain vital in any research project where collaboration of the data subject is imposed by methodological considerations.

Processing of data under a necessity framework also implies special obligations to safeguard the rights and interest of data subjects, in particular, the use of technical measures to ensure respect for the principle of data minimisation, including the default use of either pseudonymisation or complete anonymisation if the research proposed can be achieved in that manner.[69] All rights of data subjects and respective duties imposed on controllers and processors are to be observed, including specific national limitations on the processing of health and genetic data,[70] unless a derogation from data protection rights is established either by EU or national law.[71] Concerning genetic, biometric and health data, Member States are given additional

[67] Data sharing and repurposing data is a very important issue for biobanking. See below Sect. 4.4.

[68] Article 9(2)(j) GDPR.

[69] Article 89(1) GDPR.

[70] Article 9(4) GDPR.

[71] Article 89(2) GDPR.

room for manoeuvre and are allowed to introduce more stringent requirements and impose further obligations on data controllers and processors which may amount to further limitations on the processing of these special categories of data.

Article 89 GDPR gives Member States additional leeway to enact specifications and derogations from the rights of data subjects when lawfulness is based on a necessity framework.[72] Exemptions to the duties of controllers and processors may be provided in national law concerning the information requirements[73] and rights to rectification,[74] to erasure,[75] to restriction of processing,[76] to data portability[77] and to object when processing personal data 'for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes'.[78] These derogations from the rights of data subjects have a subsidiary nature and are only admissible as far as the data subject rights render impossible or seriously impair the achievement of the 'scientific or historical research purposes or statistical purposes'.[79] Derogations also have to be specified and accompanied by appropriate safeguards as to the general principles of data protection. In particular, all exemptions must follow data minimisation, proportionality and necessity principles.[80]

Biobanks may be able to use these exemptions in national law. However, the question of applicable jurisdiction has to be carefully considered, in particular the possibility that a data set might include individual data which are subject to different national exemptions and complementary rules concerning, for example, the use of genetic data.[81] If the legal basis for lawfulness is necessity for research under Article 89(2) GDPR, exemptions to the duties of controllers and processors may be provided in national or EU law concerning: (1) the right of any person to obtain from the controller confirmation as to whether or not their personal data are being processed, and the right to information concerning such processing;[82] (2) the right to rectification;[83] (3) the right to restrict processing;[84] and (4) the right to object to processing.[85] Where biobanks serve as repositories and data processing is justified for archiving purposes in the public interest under Article 89(3) GDPR, exemptions

[72] See Chapter Duguet A-M, Herveg J 'Safeguards and derogations relating to processing for scientific research: Article 89 analysis for biobank research'.

[73] Article 15 GDPR.

[74] Article 16 GDPR.

[75] Article 17 GDPR.

[76] Article 18 GDPR.

[77] Article 20 GDPR.

[78] Article 21 GDPR.

[79] Article 89(1) GDPR.

[80] Recital 156 GDPR.

[81] For an overview of existing national legislation see: Penasa et al. (2018); p. 252; Briceño Moraia et al. (2014).

[82] Article 15 GDPR.

[83] Article 16 GDPR.

[84] Article 18 GDPR.

[85] Article 21 GDPR.

established in EU or national law may also extend to the controller's obligation to notify any restriction or erasure of personal data to each third party to whom the data has been disclosed[86] and the data subject's right to data portability.[87]

The right of data subjects to request erasure of their personal data cannot be subject to national derogations under Article 89 GDPR. However, Article 17 GDPR does exempt data processing activities for archiving purposes 'in the public interest, scientific or historical research purposes or statistical purposes' in accordance with Article 89(1) GDPR[88] provided that erasing the data is likely to render impossible or seriously impair the achievement of these objectives.[89] If the data are essential but can be fully anonymised, then such an option should prevail. Controllers are under an obligation to justify the refusal to erase and to disclose information about the specific use of the data in a specific project.

Consent-Based Model

When a necessity-based lawfulness basis cannot be established, biobanks will need to resort to a consent-based model in order to avoid data protection liability. It is also a solid strategy through which to build trust and ensure recruitment of research participants while fostering the willingness of participants to provide accurate data, be monitored over time and provide multiple samples and data entries and allow multi-purpose processing.

The literature shows that prior to the GDPR Member States had different frameworks for consent.[90] Taking into account the GDPR flexibilities, the situation is likely to be maintained, at least insofar as additional specific requirements and regulatory oversight are concerned. Under the GDPR, the type of consent necessary for data processing is defined as necessarily being freely given, purpose specific, informed and unambiguous.[91] In order to be legally binding, consent does not need to be given in the form of a signed written document but should be given by a clear affirmative act. Documented oral statements and electronic means are allowed but controllers should avoid 'silence, pre-ticked boxes or inactivity' since only affirmative consent is legally binding.[92]

Compliance with the principle of fairness and transparency imposes that preformulated consent forms must be written in a manner that is intelligible and easily accessible to the data subject using clear and plain language.[93] The use of legal or

[86] Article 19 GDPR.

[87] Article 20 GDPR.

[88] Article 17(3)(d).

[89] Article 17(3)(d).

[90] Kaye et al. (2016), pp. 195–200.

[91] Article 4 11) GDPR.

[92] Recital 32 GDPR.

[93] Recital 42 GDPR.

technical terms should be avoided and, if applicable, translated into the native language of the data subjects. The standard for consent is 'free and informed consent'. Documents or information provided orally should contain clear mention of the identity of the controllers and the purpose of the data processing. Consent will not be valid if the data subject has no genuine or free choice or if refusal or withdrawal of consent is detrimental to the data subject.[94] This would be the case for multipurpose consent without the possibility to separately consent to different processing purposes or if broad consent is demanded for access to treatment or a service and the data processing exceeds what is necessary for fulfilling such goals (e.g. deposit and conservation of biological materials for future use: blood, stem cells, ova, sperm, embryos, etc.).[95]

Often in biobanking activities samples and information originate from outside the EU. In some cases, local cultural and legal traditions may result in different frameworks, rules and procedures for consent.[96] EU data protection rules are based on the EU Charter right to data protection[97] and have an extensive territorial application. Thus, if the controller or processor is established in the EU, reliance on local law or customary social norms is not possible and individual data subject informed specific consent or another legal ground for data processing remains necessary under the GDPR.

Consent should also be specific and cover every purpose and all processing activities carried out for each purpose.[98] The legislators acknowledged that in the case of data used for scientific research it is often difficult to identify beforehand all possible data processing purposes and so this opened the door to broad consent. In this sense, Recital 33 clarifies that broad consent—defined by reference to certain areas of scientific research—can be accepted if procedures comply with 'recognised ethical standards for scientific research', for example, through an ethical board review.[99] WP29 pointed out that Recital 33 does not necessarily mean that specific consent is not necessary but rather that as an exception and if research purposes cannot be specified at the time of data processing (sample collecting), it is possible to obtain valid consent and only describe the purpose in a more general manner. However, it also alerts us to the fact that 'when special categories of data are processed on the basis of explicit consent, applying the flexible approach of Recital 33 will be subject to a stricter interpretation and requires a high degree of scrutiny'.[100]

[94] Recital 42 and 43 GDPR.

[95] In such cases, specific national legislation may contain more strict rules.

[96] For an overview see for example: De Vries et al. (2017).

[97] Article 8 EU Charter.

[98] Recital 32 GDPR.

[99] See Marelli and Testa (2019), pp. 496–498.

[100] Article 29 Working Party, Guidelines on Consent under Regulation 2016/679, last Revised and Adopted on 10 April 2018.

The notion of dynamic consent[101] is indirectly accepted. On the one side, data subjects have several rights that can be exercised over a period of time: right to rectification of inaccurate personal data and to add supplementary information to incomplete data;[102] right to erasure;[103] right to restriction of processing;[104] right to data portability;[105] and right to not be subject to a decision based solely on automated processing.[106] On the other side, the re-purposing of data will require informing the data subject and renewed consent. Dynamic consent models offer biobanks the possibility to allow data subjects to exercise their rights to object to specific types of data processing, specific purposes, projects or users while simultaneously maintaining consent to a broad range of processing activities. These also simplify procedures for consent for further processing purposes and improve fairness and transparency of data processing. However, it should be noted that overall repurposing of data in biobanking remains a complex matter subject to specific national regulations[107] and where determining if the new use is compatible with the consent provided may not be easy to ascertain.[108]

Biobanks as data controllers have a duty to implement technical measures to assure that data subjects can, on request, receive the personal data provided in a structured, commonly used and machine-readable format and transmit those data to another controller.[109] It is debatable whether data portability duties apply only to raw data or also to established correlations, probabilities or predictions, for example, a diagnosis. As long as a person is identifiable then the information is considered personal data and thus subject to the GDPR.[110] Inferred data and derived data, such as the outcome of an assessment regarding the health of a user, are, according to WP29, excluded from the right to data portability.[111] Furthermore, this information

[101] Kaye et al. (2015), pp. 141–146.

[102] Article 16 GDPR.

[103] Article 17 GDPR.

[104] Article 18 GDPR.

[105] Article 20 GDPR.

[106] Article 22 GDPR.

[107] See: Tassé (2016), pp. 207–216; Kondylakis et al. (2017), pp. 282–292.

[108] See the landmark Italian case concerning the acquisition by United Kingdom–based commercial company Tiziana Life Sciences Plc of Shardna an Italian genomic biobank (Tribunal of Cagliari, Sentenza n. 1569, 6 June 2017) described in Marelli & Testa n.101; see also recent Clinical Research Development Ireland (CRDI) 'Submission to the Data Protection Commission on the topic of the General Data Protection Regulation in relation to Biobanking'(3 May 2018), signed by 28 Representatives of Irish Research Institutions. Available: https://www.crdi.ie/wp-content/uploads/2018/06/CRDI_Submission_GDPR-and-Biobanking.pdf.

[109] Article 20 GDPR.

[110] Article 4 (1) GDPR defines an identifiable person as 'one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'.

[111] Article 29 data protection working party, Guidelines on the Right to data Portability, adopted on 13 December 2016.

may constitute a trade secret or be copyright protected and proportionality arguments may arise, while specific contractual or patient rights statutory provisions may provide further obligations.

In the context of big data analytics, where data are obtained from a plurality of sources, the controller always has general information duties that may be difficult to comply with, including providing individual information concerning categories of data, origin, legal basis and purpose of processing and use in automated decision-making.[112] These duties are waived if providing information to data subjects proves impossible or involves a disproportionate effort, and where the processing is for scientific research purposes and compliance with such duties would render impossible or seriously impair the research.[113] Either way, repurposing of data must always have a legal basis; either it has to be covered by original consent or an exception.

Consent can be withdrawn and the data subject can request that further processing is restricted or that the data is erased. The right to erasure, known as the right to be forgotten, is often considered a potential challenge. However, research activities are protected if the data are necessary for research and their erasure would 'render impossible or seriously impair the achievement of the objectives of that processing'.[114] This is not a complete exemption; an erasure request must still be complied with if under the specific circumstances that individual's personal information is not essential and can be erased without compromising the entire study. In any case, if the data are not erased due to being essential, it might have to be erased from other research projects and cannot continue to be processed in the future unless another ground for processing exists.

Finally, consent to participation in scientific research activities in clinical trials is subject to specific legislation—the Clinical Trials Regulation (CTR).[115] GDPR principles and other rules remain applicable to data processing in the context of clinical trials.[116] Consent for data processing in the context of biobanking samples and data originated or procured for clinical trials will also follow the GDPR rules and should not be confused with informed consent for participation in clinical trials and/or medical treatment.[117] Informed consent for these activities is regulated by specific frameworks and follows a different legal reasoning.[118] As explained by the EDPB in Opinion 3/2019, the provisions on informed consent in the Clinical Trials

[112] Article 14(1) GDPR.

[113] Article 14(5) GDPR.

[114] Article 17(3)(d) GDPR.

[115] Articles 28–35, Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC, *OJ L 158, 27.5.2014, p. 1–76.*

[116] See the recent, European Data protection board, Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR) (art. 70.1.b)) Adopted on 23 January 2019.

[117] Idem, para 15.

[118] Minssen and Rajam (2019); Chico (2018), p. 116.

Regulation[119] respond primarily to core ethical requirements of research involving humans subjects and derive from the Helsinki Declaration. The obligation to obtain informed consent of participants in a clinical trial is primarily required to ensure respect for the right to human dignity and the right to integrity of individuals under Articles 1 and 3 of the Charter of Fundamental Rights of the EU and is not an instrument for data protection compliance.[120]

This means that informed consent obtained for clinical trials may not be sufficient for data processing purposes. In particular, a 'clear situation of imbalance of powers between the participant and the sponsor/investigator will imply that the consent is not 'freely given' in the meaning of the GDPR'[121] (e.g. when a participant is not in good health, belongs to an economically or socially disadvantaged group or is in any situation of institutional or hierarchical dependence). Therefore, consent will not be the appropriate legal basis in most cases and other legal bases than consent must be relied upon.[122] Biobanks storing samples or data obtained or used in clinical trials have to conduct a separate assessment on the legal basis of data processing to rely upon and eventually obtain consent for initial or further biobanking activities, unless the so-called presumption of compatibility provided under Article 5(1)(b) GDPR can be relied upon under the specific circumstances.[123]

4.3 Fairness and Transparency of Data Processing

Although biobanks operating under the framework for lawfulness established under Article 89 'Interest for scientific research-based model' are exempted from a number of specific obligations, the principle of transparency imposes an obligation to inform data subjects at the time data are obtained of the following: (1) identity and the contact details of the controller and, where applicable, of the controller's representative; (2) contact details of the DPO; (3) purposes and legal basis of the processing; (4) recipients or categories of recipients of the personal data; and (5) whether the controller intends to transfer personal data to a third country or international organisation, and the existence or absence of an adequacy decision by the Commission, or reference to the appropriate or suitable safeguards and the means by which to obtain a copy of them or where they have been made available.[124]

In addition to this information, biobanks acting as data controllers also have a duty to provide to the data subject at the time personal data are obtained additional information to ensure fair and transparent processing, namely, (1) length of time

[119] CTR Chapter V, Article 28 e sq.

[120] EDPB Opinion 3/2019, para 16.

[121] EDPB Opinion 3/2019, para 20.

[122] Idem.

[123] EDPB Opinion 3/2019, para 29-32.

[124] Article 13(1) GDPR.

data will be stored (either a fixed date or criteria used to determine it); (2) details about the right to lodge a complaint with a supervisory authority; and (3) the existence of automated decision-making, including profiling and meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.[125]

Although the rule is that data subjects have the right to object to automated decision-making and profiling, automated decisions and profiling (e.g. diagnostic, epidemiologic studies, categorisations of genetic risk, etc.) based on special categories of data, such as health and genetic data, are not prohibited. In fact, these can be acceptable if based on explicit consent for specified purposes or if based on the necessity of the processing for reasons of substantial public interest.[126]

If the ground for data processing is consent, biobanks as data controllers are also required to provide information on the existence of the right to request access to and rectification or erasure of personal data or restriction of processing concerning the data subject or to object to processing as well as the right to data portability. Biobanks will also be obliged to inform data subjects that they have the right to withdraw consent at any time, and that this will not affect retroactively the lawfulness of previous processing. These obligations will not subsist if data is processed based on other grounds.[127]

4.4 Purpose Limitation of Data Processing

Data sharing is increasingly necessary for scientific research, and there is a growing international trend towards open science,[128] with major funding agencies and scientific journals imposing data sharing policies.[129] Such policies may implicitly result in imposing the need to share or make public available research data outside the EU. In their turn, EU initiatives also place considerable emphasis on open research data and open access to scholarly publication and communication and reuse of scientific information.[130]

[125] Article 13(2) GDPR

[126] Article 13(2)(f), Article 22(4) and Article 9(2 (a) and (g) GDPR.

[127] Article 13(2) GDPR.

[128] Groves and Godlee (2012), p. e4383.

[129] Taichman et al. (2017), pp. 63–65; National Institutes of Health (NIH) (2003); Wellcome Trust (2017); European Commission DG for research and Innovation (2017).

[130] Commission Recommendation of 17 July 2012 on access to and preservation of scientific information (2012/417/EU); see also Declaration of the Budapest Open Access Initiative https://www.budapestopenaccessinitiative.org/read; Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities https://openaccess.mpg.de/67605/berlin_declaration_engl.pdf; The ECHO Charter https://echo.mpiwg-berlin.mpg.de/policy/oa_basics/charter, and the Bethesda Statement on Open Access Publishing http://legacy.earlham.edu/~peters/fos/bethesda.htm.

Biobanking research by its nature involves the possibility to re-use and repurpose collected samples and information in several research projects. New digital technologies offer increased possibilities to cross-reference large quantities and types of data from multiple sources. The interpretation of the principle of purpose limitation has become a central issue in biobanking as both data sharing and data repurposing raise considerable data protection and ethical issues;[131] a balance needs to be achieved with the protection of the rights of data subjects.

The principle of purpose limitation ensures that as a rule all data must be 'collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes',[132] and it can be particularly controversial to apply in the context of biobanking sharing and re-use of research data. Subsequent uses may rely either on consent or another ground for lawfulness; both these grounds have to be established at the time a biological sample, tissue or information is collected and further processing has to be compatible with the purpose for which the personal data are initially collected.[133] If the lawfulness of data processing is based on necessity for archiving purposes in the public interest, scientific or historical research, the re-purpose of data for archiving or research is generally presumed compatible with the original purpose as long as the controller demonstrates respect for the individual rights and freedoms of the data subject and implements appropriate safeguards, such as pseudonymisation (unless this is impossible or impairs the archiving or research purposes).[134] However, the presumption appears to only apply if it is the same type of research or research project, for example, the EDPB does not think that it necessarily applies to clinical trials data reuse.[135] Moreover, if the data processing is based on another lawfulness ground, then compatibility can never be presumed and it is either necessary to establish that the specific research conducted is compatible with the original purpose or predict and establish at the time of data collection several possible specific, explicit and legitimate data purposes.

When biobanks intend to further process the personal data for a purpose other than that for which the personal data were collected, information must be provided to the data subjects prior to that further processing concerning such further processing and its purpose, as well as any other relevant information.[136] Moreover, often biobanks will store and process data that was not obtained directly from data subjects but instead was originally collected from a third party, for example, biological samples obtained in a clinical setting or use of health records. In such cases, and in

[131] For an overview on open questions see: Global Forum on Bioethics in Research (2018).

[132] Article 5(1)(b) GDPR.

[133] Article 6(4) GDPR. See with adaptations Article 29 Data Protection Working Party Opinion 03/2013 on purpose limitation Adopted on 2 April 2013.

[134] Article 5(1)(b) and Article 89(1) GDPR; Recitals 157 to 160.

[135] See EDPB Opinion 3/2019 para 28, recognizing that further guidance in this respect is necessary.

[136] Article 13(3) GDPR.

the absence of more specific national or EU legislation,[137] information duties subsist in accordance with Article 14 GDPR. There are, however, some exceptions: compliance with information duties is not required if the data subject already has the information. Regarding processing based in public interest and research purposes, there is no duty to provide information if this has been proven to be impossible or would involve a disproportionate effort, or if it is likely to render impossible or seriously impair the objectives of the biobanking activity. The biobank neverthelesss must take appropriate measures to protect the data subjects' rights and freedoms and legitimate interests, including making the information publicly available.[138]

4.5 Data Protection by Design

As controllers, biobanks are also responsible to implement measures leading to 'data protection by design and default'. Data protection by design implements the principle of data minimisation and is imposed under a standard of reasonability taking into consideration a number of factors, such as the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing.[139] Appropriate technical measures include pseudonymisation[140] but also measures for ensuring that personal data are only used if necessary for a specific purpose. This means that all data processed must be relevant for a specific research question. The data minimisation obligation also applies to ensure that the amount of personal data collected, the extent of their processing, the period of their storage and who is granted access is linked and necessary for the purpose of data processing.[141] Generally, biobanks acting as data controllers are always responsible for implementing appropriate technical and organisational measures to ensure compliance with data protection rules. Compliance may be demonstrated inter alia by specific data protection policies, adherence to approved codes of conduct[142] or through use of approved certification mechanisms.[143]

[137] Article 14(5)(c) and (d) GDPR 'obtaining or disclosure is expressly laid down by Union or Member State law to which the controller is subject and which provides appropriate measures to protect the data subject's legitimate interests; or where the personal data must remain confidential subject to an obligation of professional secrecy regulated by Union or Member State law, including a statutory obligation of secrecy.'

[138] Article 14(5)(a) and (b) GDPR.

[139] Article 25(1) GDPR.

[140] Idem. Cf. notion of pseudo-anonymisation in Article 4(5) GDPR with different understandings in other normative sources see: Phillips et al. (2017), pp. 483–496.

[141] Article 25(2) GDPR. On compliance strategies See: Holub et al. (2018), pp. 97–105.

[142] Article 24 and 40 GDPR.

[143] Articles 24 and 42 GDPR.

4.6 Data Stewardship

Biobanks are also entrusted with data stewardship duties. These are formulated as the principles of data accuracy, storage limitation, integrity and confidentiality. Data controllers have the obligation to keep records of all processing activities. This obligation is related to the principle of transparency and has the purpose of guaranteeing compliance with data subjects' rights and preventing controllers from alleging insufficient knowledge based on deficient records as a defence.[144] Biobanks acting as controllers are also responsible for guaranteeing the security of data processing activities,[145] cooperating with data protection authorities (DPA);[146] and notifying the DPA of any data breaches within 72 h[147] and each data subject provided that there is a high risk to their rights and freedoms. Data controllers should conduct data protection impact assessments (DPIAs),[148] implement measures to mitigate the risks discovered and consult with data protection authorities where such DPIAs determine a high risk that cannot be mitigated.[149] Biobanks process special categories of personal data and therefore DPIAs are mandatory.[150] Controllers and processors may also be responsible for jointly designating a DPO.[151] This duty will apply to biobanks and biobank researchers insofar as their core activity entails processing large amounts of special categories of personal data.[152]

5 Conclusion

The recent reform of data protection rules in the EU is in several ways a positive step in the direction of balancing individual rights and ensuring that scientific research and innovation in a data-driven economy are not hindered. A number of exemptions and exceptions are provided for research activities, with Article 89 GDPR making it possible for Member States to adopt further exceptions and exemptions. While this has a positive side, it also favours forum shopping, creates difficulties in pan-European studies and risks reducing harmonisation and transforming the GDPR almost into a de facto *directive* as far as the scientific research context is concerned.

[144] Article 30 GDPR.
[145] Article 32 GDPR.
[146] Article 31 GDPR.
[147] Article 33 GDPR.
[148] Article 35 GDPR.
[149] Article 36 GDPR.
[150] Article 35 (3) (b) GDPR.
[151] Article 37 GDPR.
[152] Article 37(1)(c) GDPR.

Its broad scope of geographic application expands the application of GDPR to many data processing situations that have a connection with the EU even when the data are not processed in the EU, i.e. either through the data controller or data processor being considered established in the EU or when the data pertain to data subjects in the EU. Local data protection rules might no longer be considered sufficient and, given the level of international collaboration in the field of biobanking, the GDPR rules might become a *de facto* international data protection standard.

The main restriction imposed on data controllers and processors is the duty to ensure the lawfulness of such activities. The GDPR contains two main legal bases for data processing of interest to biobanks: consent-based model and necessity-based model. It will remain critical to carefully consider which to apply to each data set because combining data sets based on different lawfulness grounds may generate increased compliance complexity.

Finally, the GDPR maintains a regulatory approach based on types of data (personal and special) and general lawfulness grounds for processing. It does not provide specific rules for particular activities of data processing and types of data uses. Legal persons data are left subject to national laws as the GDPR rules only applies to natural persons data and there is no differentiation between types of more or less intrusive uses. It does not clearly differentiate between raw data and inferred data and derived data. Neither does it consider the privacy impact of cumulative or network effects of data aggregation and cross-reference.

Compliance with the GDPR presents challenges for biobank and biobank researchers using advanced digital technologies. The use of big data analytics has brought tremendous benefits to scientific research, particularly in the field of genetics. Developments in this area include cost-effective sequencing of entire genomes and the possibility to share and combine multiple sources of complementary data. The very nature of research using big data analytics in general and genetic data in particular suggests that compliance may be onerous and difficult to implement in research protocols and institutional procedures. As we move deeper into a digitalised and data-driven society, particularly problematic data uses will require further clarification and improved approaches to data protection. Growing use of AI and big data analytics in biobanking activities means that special attention to compliance procedures will be necessary and that in the long term further legal developments and interpretative guidance should be expected.

References

Article 29 Data Protection Working Party, Guidelines on Consent under Regulation 2016/679, last Revised and Adopted on 10 April 2018

Article 29 Data Protection Working Party, Guidelines on the Right to data Portability, adopted on 13 December 2016

Article 29 Data Protection Working Party, Opinion 03/2013 on purpose limitation Adopted on 2 April 2013

Article 29 Data Protection Working Party, Opinion 1/2010 on the concepts of "controller" and "processor", adopted on 16 February 2010

Beier K, Lenk C (2015) Biobanking strategies and regulative approaches in the EU: recent perspectives. J Biorepository Sci Appl Med 3:69–81

Briceño Moraia L et al (2014) A comparative analysis of the requirements for the use of data in biobanks based in Finland, Germany, the Netherlands, Norway and the United Kingdom. Med Law Int 14(4):187–212

Chico V (2018) The impact of the general data protection regulation on health research. Br Med Bull 128(1):109–118, p.116

Commission Recommendation of 17 July 2012 on access to and preservation of scientific information (2012/417/EU)

Council of Europe Recommendation CM/Rec 2016 (6) of the Committee of Ministers to member states, adopted by the Committee of Ministers on 11 May 2016 at the 1256th meeting of the Ministers' Deputies

Council of Europe Recommendation Rec 2006 (4) of the Committee of Ministers to member states on research on biological materials of human origin, adopted by the Committee of Ministers on 15 March 2006 at the 958th meeting of the Ministers' Deputies

Council of Europe's Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Coe ETS No.108), Strasbourg, 28/01/1981

De Vries J et al (2017) Regulation of genomic and biobanking research in Africa: a content analysis of ethics guidelines, policies and procedures from 22 African countries. BMC Med Ethics 18(8):8

Declaration of the Budapest Open Access Initiative. https://www.budapestopenaccessinitiative.org/read; Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Available at: https://openaccess.mpg.de/67605/berlin_declaration_engl.pdf

Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use, OJ L 121, 1.5.2001, p. 34

Directive 2004/23/EC of the European Parliament and of the Council of 31 March 2004 on setting standards of quality and safety for the donation, procurement, testing, processing, preservation, storage and distribution of human tissues and cells, *OJ L 102, 7.4.2004, p. 48–58*

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, *Official Journal L 281, 23/11/1995 P. 0031 – 0050*

Dove ES (2019) The EU general data protection regulation: implications for international scientific research in the digital era. J Law Med Ethics 46(4):1013–1030. https://doi.org/10.1177/1073110518822003

ECHO Charter. https://echo.mpiwg-berlin.mpg.de/policy/oa_basics/charter, and the Bethesda Statement on Open Access Publishing http://legacy.earlham.edu/~peters/fos/bethesda.htm

European Commission DG for research and Innovation (2017) H2020 programme guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020. Available at: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

European Data protection board, Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR) (art. 70.1.b)) Adopted on 23 January 2019

Fransson MN, Rial-Sebbag E, Brochhausen M et al (2015) Toward a common language for biobanking. Eur J Hum Genet 23(1):22–28

Global Forum on Bioethics in Research (2018) Background paper: the ethics of data sharing and biobanking in health research, Meeting in Cape Town, 13–14 November 2018. Available at: http://www.gfbr.global/wp-content/uploads/2018/04/GFBR-2018-background-paper-FINAL.pdf

Groves T, Godlee F (2012) Open science and reproducible research. BMJ 344:e4383
Hellstadius A, Schovsbo J (2018) You told me, right? - Free and informed consent in European patent law. In: Minssen T, Herrmann JR, Schovsbo J (eds) Global genes, local concerns: legal, ethical and scientific challenges in international biobanking. Edward Elgar, Cheltenham
Hewitt R, Watson P (2013) Defining biobank. Biopreserv Biobank 11(5):309–315
Holub P et al (2018) Enhancing reuse of data and biological material in medical research: from FAIR to FAIR-Health. Biopreserv Biobank 16(2):97–105
Judgment of the Court (Grand Chamber), 13 May 2014, *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos* (AEPD) and Mario Costeja González, Case C-131/12, *Digital reports:* ECLI identifier: ECLI:EU:C:2014:317
Judgment of the Court (Grand Chamber) of 5 June 2018, *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH*, Case C-210/16, OJ C 260, 18.7.2016, ECLI:EUC:2018:388
Kaye J et al (2015) Dynamic consent: a patient interface for twenty-first century research networks. Eur J Hum Genet 23(2):141–146
Kaye J et al (2016) Consent for biobanking: the legal frameworks of countries in the BioSHaRE-EU Project. Biopreserv Biobank 14(3):195–200
Kondylakis H et al (2017) Donor's support tool: enabling informed secondary use of patient's biomaterial and personal data. Int J Med Inform 97:282–292
Marelli L, Testa G (2019) Scrutinizing the EU general data protection regulation. Science 360(6388):496–498
Minssen T, Rajam BM (2019) Clinical trials data transparency & GDPR compliance: what are the effects on data sharing and open innovation?, Science and public policy scaa014
Morrison M et al (2017) The European general data protection regulation: challenges and considerations for iPSC researchers and biobanks. Regen Med 12(6):693–703
National Institutes of Health (NIH) (2003) Final NIH statement on sharing research data. Available at: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html
Nicola P (2015) Spanish regulation on biobanks. J Law Med Ethics 43(4):800–815
Nordberg A, Minssen T (2016) A 'Ray of Hope' for European stem cell patents or 'Out of the Smog into the Fog'?: an analysis of recent European case law and how it compares to the US. IIC 47(2):138–177
OECD, Recommendation on human biobanks and genetic research databases, adopted by the OECD Council on 22 October 2009
Penasa S et al (2018) The EU general data protection regulation: how will it impact the regulation of research biobanks? Setting the legal frame in the Mediterranean and Eastern European area. Med Law Intern 18(4):241–255
Phillips A, Borry P, Shabani M (2017) Research ethics review for the use of anonymized samples and data: a systematic review of normative documents. Account Res 24(8):483–496
Protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (CoE ETS No. 223), 128th Session of the Committee of Ministers, Elsinore, 17–18 May 2018
Regulation 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC, OJ L 158 27.05.2004, p. 1–76
Sandor J, Drakopoulou A, Bard P (2009) The legal regulation of biobanks: National Report: Greece. CELAB Paper Series No. 2. Available at https://doi.org/10.2139/ssrn.2295967
Shaw DM, Elger BS, Colledge F (2014) What is a biobank? Differing definitions among biobank stakeholders. Clin Genet 85(3):223–227
Taichman DB et al (2017) Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. Ann Intern Med 167(1):63–65
Tassé AM (2016) A comparative analysis of the legal and bioethical frameworks governing the secondary use of data for research purposes. Biopreserv Biobank 14(3):207–216

Watson PH (2014) Biobank classification: communicating biorepository diversity. Biopreserv Biobank 12(3):163–164

Wellcome Trust (2017) Policy on data, software and material management and sharing at https://wellcome.ac.uk/funding/managing-grant/policy-data-software-materials-management-andsharing