The Impact of the GDPR on the Governance of Biobank Research

Mahsa Shabani, Gauthier Chassang, and Luca Marelli

Abstract Governance of health and genomic data access in the context of biobanking is of salient importance in implementing the EU General Data Protection Regulation (GDPR). Various components of data access governance could be considered as 'organizational measures' which are stressed in the Article 89(1) GDPR together with technical measures that should be used in order to safeguard rights of the data subjects when processing data under research exemption rules. In this chapter, we address the core elements regarding governance of biobanks in the view of GDPR, including conditions for processing personal data, data access models, oversight bodies and data access agreements. We conclude by highlighting the importance of guidelines and policy documents in helping the biobanks in improving the data access governance. In addition, we stress that it is important to ensure the existing and emerging oversight bodies are equipped with adequate expertise regarding using and sharing health and genomic data and are aware of the associated informational risks.

M. Shabani (✉)
Metamedica, Faculty of Law and Criminology, Ghent University, Ghent, Belgium
e-mail: mahsa.shabani@ugent.be

G. Chassang
Inserm, Faculté de Médecine, Toulouse, France
e-mail: gauthier.chassang@bbmri-eric.eu

L. Marelli
Life Sciences and Society Lab, Centre for Sociological Research (CeSO), University of Leuven, Leuven, Belgium

Department of Experimental Oncology, IEO, European Institute of Oncology IRCCS, Milan, Italy
e-mail: luca.marelli@kuleuven.be

S. Slokenberga et al. (eds.), *GDPR and Biobanking*, Law, Governance and Technology Series 43, https://doi.org/10.1007/978-3-030-49388-2_4

1 Introduction

Governance of health and genomic data access in the context of biobanking is of salient importance in implementing the EU General Data Protection Regulation (GDPR). Various components of data access governance could be considered as 'organizational measures' which are stressed in the Article 89(1) GDPR together with technical measures that should be used in order to safeguard rights of the data subjects when processing data under research exemption rules. By establishing adequate governance mechanisms from the outset in the process of personal data processing, the ultimate goal of the regulation in terms of 'privacy by design' will be facilitated, in which data protection safeguards will be built into the products and services from the earliest stage of development.

According to the GDPR Article 9(2)(j), personal data, including sensitive data, could be processed for scientific research purposes under the conditions set out in the Article 89. As Article 9(2)(j) states: 'processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.'

In principle, adopting adequate governance models that are foreseen by the GDPR will establish additional controls, to protect the rights of the data subjects when processing personal data for research purposes. A similar approach has been supported by a report on the *Collection, linking and use of data in biomedical research and health care* by Nuffield Council on Bioethics, which noted, 'Because of the risk of misuse and consequential privacy infringement, de-identification and consent measures may be supplemented by further governance arrangements.'[1]

One key element in biobank governance is developing transparent and fair data access rules, which should address the core elements regarding data access review and oversight procedures. Generally speaking, rules for data access should delineate criteria for data user's qualification, the review procedure, and terms and conditions of access. The ultimate goal is to decrease the risks of harms to the research participants that may arise from unauthorized access to the datasets for unintended purposes. In principle, the development of the data sharing and access rules must be in compliance with the applicable national laws. The relevant international and national data sharing policies and guidelines that are issued by various professional communities may guide the development of data access rules.

Moreover, data access rules should be developed in the view of suitable data access models, which could range from fully open-access to controlled-access. The nature of the data in terms of identifiability and the associated privacy risks for the data subjects significantly influences the model of data access. It should be noted that biobanks and data-intensive genomics and health studies might use external

[1] Nuffield Council on Bioethics (2014), p. 7.

data repositories for data sharing such as the NIH database of Genotypes and Phenotypes (dbGaP) or the European Genome-phenome Archive (EGA).[2] This could be requested by funding organizations or journals in order to facilitate broad access to the data. In case of using external databases, it is essential for the researchers to ensure that the data governance models of the databases conform with the applicable national laws and institutional policies.[3]

In this chapter, we address a number of issues essential in discussion regarding governance of biobanks in the view of GDPR. First, we will investigate the GDPR's relevant provisions regarding processing personal data under research exemption. This is particularly pertinent for the governance of biobanks, as personal data harvested from biological samples may include a wide range of health and genomic data. Second, we will provide an overview of the major data access models, namely open access, registered access and controlled access. This overview will enable us to show the level of control that biobanks could maintain on data based on the selected model of data access. Finally, we will review the functions of the relevant oversight committees in the framework of governance of data access. Some of these oversight committees, such as Data Access Committees are not defined by the GDPR, yet they are essential in the governance of data access in biobanks. We will also refer to data transfer agreements as an important tool used in the governance of data access.

2 Processing Personal Data for Scientific Research Purposes

The GDPR provides a certain degree of flexibility for the processing of personal data for scientific research purposes. Notably, the GDPR upholds a 'research exemption' to the general prohibition otherwise imposed on the processing of 'special categories of data'[4] (a label under which are grouped sensitive data like genetic, biometric and health-related data that are recognized as warranting the implementation of higher forms of protection from the part of data controllers.[5]) In addition, Article 6 recognizes processing personal data for public interest or legitimate interest in the list of lawful grounds for processing data. When read in conjunction with Art. 9(2)(j), this can, in turn, provide a legal basis for processing data for scientific research purposes. The so-called research exemption allows the processing of data for scientific research purposes, where the processing is proportionate to the aim pursued, that is, only personal data which is adequate and relevant for the purposes of the processing is collected and processed.

[2] Paltoo et al. (2014), pp. 692–695.

[3] Mascalzoni et al. (2019).

[4] Article 9(2)(j), GDPR.

[5] Recital 53, GDPR.

Additionally, the Regulation[6] relaxes the stringent requirements for specific consent and data storage—two key aspects directly impinging on biobanking—, allowing use of broad consent whenever required by the intended research purposes,[7] and to extend the period in which personal data can be legally stored.[8,9]

Crucially, subject to the provision of technical and organizational safeguards, the GDPR further allows Member States to introduce derogations from the core data subject rights of data access, rectification, restriction of processing, and object to the processing,[10] whenever upholding such rights is 'likely to render impossible or seriously impair' the achievement of the desired scientific research purposes, and such derogations are deemed essential for the fulfilment of these purposes.[11] More in general, in line with the principle of subsidiarity and the (historically) national competence in the field of health, Article 9(4) of the Regulation allows Member States to maintain or introduce further conditions, including limitations, with regard to the processing of genetic, biometric and health-related data. On a par with the derogations foreseen under Article 89(2) that are further elaborated in this volume by Anne-Marie Duguet and Jean Heveg, this could potentially lead to the fragmentation of the regulatory landscape underpinning the operations of European biobanks.[12]

3 Pseudonymized and Anonymized Data

3.1 Introductory Remarks

In order to identify the adequate organizational and technical measures in accessing and sharing genomic and health data in the context of biobanks, it is crucial to investigate the status of data, and whether the data is being considered as personal data under the GDPR. A relevant distinction enshrined in the GDPR, with significant implications for the processing and governance of access to sensitive data in the field of biobanking, is the one between pseudonymized and anonymized data.

[6] Recital 33, GDPR.

[7] Article 29 Working Group Party (2018).

[8] Article 5(1)(e), GDPR.

[9] Marelli and Testa (2018), pp. 496–498.

[10] Article 15, 16, 18 and 21, GDPR.

[11] Article 89(2), GDPR. LERU (2016).

[12] For insights in how Article 89(2) has been implemented in different EU Member States and EEA states, see Tzortzatou et al. 'Biobanking across Europe post-GDPR: A deliberately created fragmented landscape' in this volume.

3.2 Pseudonymized Data

Pseudonymized data are defined, in Article 4(5), as data that 'can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person'. This is typically the case of key-coded data, which allows (among other things) the traceability and correlation of genotypic and phenotypic data, as well as the possibility to recontact research participants, while still preserving the de-identification of personal data in the day-to-day operations of the organization. Accordingly, *insofar* as they are not irreversibly de-identified, pseudonymized data are considered as personal data, falling under the scope of the GDPR.

On the contrary, according to Recital 26, *irreversible* de-identification is defined as 'information which does not relate to an identified or identifiable natural person' or as 'personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable'. As further specified in Recital 26, anonymized data fall outside the remit of the GDPR. However, it should be noted that the act of anonymization itself should be considered as an act of processing personal data, which should occur, accordingly, in compliance with the GDPR.

3.3 Anonymization of Data

When we focus on anonymization, the main question to be addressed is: Under what circumstances, if any, can genomic and health data be anonymous in light of the GDPR?[13] Interestingly, the GDPR differs conspicuously, in this respect, from other major data protection legislations, such as the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule in the US.[14] Within the Privacy Rule, the Safe Harbor standard for achieving the de-identification of personal data singles out 18 distinct identifiers, the removal of which is said to make the resulting information 'not individually identifiable', and thus anonymous.[15]

Differently from this approach, recital 26 of the GDPR states instead that personal data should be considered anonymous insofar as the data subject cannot be identified 'by any means reasonably likely to be used [...] either by the controller or by any other person'.[16] To ascertain whether means are reasonably likely to be used to identify the natural person, the GDPR further states that 'account should be taken

[13] For a broader overview of this issue in relation to genomic data, cf. Shabani and Marelli (2019).

[14] Shabani et al. (2018).

[15] U.S. Department of Health & Human Services (2012), p. 6.

[16] Recital 26, GDPR; see also: Court of Justice of the European Union (CJEU), Case C-582/14: Patrick Breyer v Bundesrepublik Deutschland, ECLI:EU:C:2016:779.

of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments' (Recital 26). In addition, opinion 05/2014 of the Article 29 Working Party has outlined other factors that should be taken into consideration, such as the existence of publicly available data which can be cross-referenced with the original dataset, thus heightening the risk of de-anonymization. As such, and in line with the overall decentralized thrust of the Regulation, the GDPR can be said to adopt a context-base criterion to determine whether personal data should be considered as irreversibly de-identified (and thus anonymous), bestowing upon controllers the responsibility to address such a question (is there a 'reasonable likelihood' that re-identification techniques can be effectively used to de-anonymize my given dataset?) in the context of their concrete processing activities.

4 Governance Models for Accessing Genomic and Health Data

4.1 Governance Models: An Overview

Samples and data collected by biobanks can be accessed for various research purposes. Such access may not be limited only to the researchers/clinicians who collected the data, but also a broader range of researchers. Adopting adequate governance models would assist to protect data subjects against potential privacy breaches. The current governance model can be grouped under three major models of open access, controlled-access and registered access, which are explained below.

4.2 Open-Access

Open-access models generally refer to making data available for the users through various online platforms without any constraint. Sharing data through open-access models has been initially pursued by the Human Genome Project, which sequenced the whole human genome for a first time in the course of 13 years.[17] However, the concerns related to identifiability of genomic data that has been demonstrated by a number of re-identification studies, questioned the adequacy of adopting such model when sharing health and genomic data.[18] Consequently, genomic data have

[17] Cook-Deegan and McGuire (2017), pp. 897–901.

[18] Homer et al. (2008), pp. 321–324.

been moved to the controlled-access databases.[19] This has been mainly the case when sharing personal level information rather than aggregate data.

A key question here is when genomic data could be considered as non-identifiable under GDPR, therefore suitable for sharing through open-access models? The regulation states that the principles of data protection 'should not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable' (Recital 26).

As it has been shown in the previous part, GDPR adopts a context-based criterion to determine whether personal data should be considered as irreversibly de-identified (and thus anonymous) and do not define the standards for de-identification itself. Hence, it is important, to decide when data can be considered as anonymous and do not fall under GDPR protection. Thereby, this is the responsibility of the data controllers to confirm whether the data is not identifiable by reasonable likelihood. For example, in the context of genomics, only sharing variant-level aggregate data may not be considered as identifying personal data, therefore adopting open-access model for sharing such data would seem acceptable under the GDPR. In a same vein, recently National Institutes of Health (NIH) updated its Genomic Data Sharing Policy and allowed unrestricted access to genomic summary results that do not raise privacy concerns.

4.3 Controlled-Access

In the view of privacy concerns when sharing health and genomic data, adopting a controlled-access model for data sharing is favored. Thereby, the data controllers can set rules for data access and limit access to the datasets to the approved users and under the determined terms and conditions. Such access control mechanisms can be considered as technical and organizational measures, which are mentioned in Article 89(1). Although there is no single model for controlled-access, a common approach is to establish oversight committees, or so-called Data Access Committees (DACs) to review the data access requests for the purpose of approval or disapproval. One of the important aspects of controlled-access data sharing is to use tools such as data access agreements (see Sect. 5), which are legally binding documents, in order to hold users accountable against potential misuses of data. This is in contrast with the open-access model where the users do not enter to any agreement with the data holders.

Oversight by DACs could be considered as an example of organizational measures that have been stressed in Article 89. Thereby, further safeguards could be offered to protect the privacy of the data subjects and ensure the downstream data

[19] Rodriguez et al. (2013), pp. 275–276.

uses conform to the original consent forms.[20] However, the recent studies have showed that the current oversight by DACs are not always efficient or effective.[21] One major reason for the identified shortcomings is DACs are not always equipped with sufficient tools and oversight mechanisms to effectively review the data access requests or detect the potential violations of data access agreements.

In response, novel approaches to data access oversight are being developed. In particular, it has been suggested to replace or supplement review by DACs by automated tools.[22] In addition, not all steps of data access review are deemed to be necessary for all types of health and genomic data sharing. In the next section, we will provide an overview of one of these recently suggested methods for data governance, namely the Registered Access model.

4.4 Registered Access

Registered access is likely to be suitable as a mechanism for access to data types that are less sensitive and low risk, such as non-stigmatizing health-related data from non-vulnerable individuals who would expect, or have consented to, data sharing for the purposes envisaged.[23] This model would focus primarily on ensuring that the data users are *bona fide* researchers. The rationale behind the registered model is that if processing data is not creating high risks of identifiability and the users are trusted, then further access review (for instance reviewing the ethical or scientific aspects of the proposals) would be redundant or disproportionate.

The 'registered access' model hinges on a number of core elements, namely authentication, authorization and attestation. First, the data use applicants should provide personal and professional information within a registration process, including their name, title, position, affiliation, email address, institutional website and mailing address for the purpose of authentication. In contrast to a controlled-access model, a registered-access model would not entail verification on a case-by-case basis by a DAC of the users' qualifications. In addition, the applicants should declare that they are 'bona fide' researchers in order to be authorized access. At last, the applicants should agree with the terms and conditions of the data access. Within the registered access model, data users would not need to sign a data access agreement in a paper-based format but could instead agree via clickwrap-type online agreements. Indeed, the procedure for signing data access agreements by DACs,

[20] Shabani and Borry (2017), pp. 149–156.

[21] Shabani and Borry (2016), pp. 892–897.

[22] Woolley et al. (2018), p. 17.

[23] Dyke et al. (2016), pp. 1676–1680.

and users and their home institution, is administratively heavy and this proposed alternative approach could reduce pressure on DACs and create rapid, open and efficient access to data.

A Registered-access model is only one proposed solution in response to the limitations of the controlled-access model. It is expected that novel governance models will emerge in the coming years in order to address the identified shortcomings of the controlled-access models, and in line with the principles of responsible data sharing. In addition to emerging governance mechanisms, novel technical solutions are also proposed,[24] including the introduction of federated networks in which multiple distributed databases are connected.[25] By using federated networks, users would be able to have (a level of) access to data in a protected virtual environment, and each database would be able to monitor data uses in real time. To date, few models of federated data computation have been suggested.[26] Considering the limitations of controlled-access models, there is a pressing need for the introduction of such innovative solutions. Concurrently, it is important to ensure the core elements of secure data computational environments are in line with data protection principles.

5 Relevant Data Sharing and Access Oversight Bodies and Tools

5.1 Data Access Committees

The need to establish an extra layer of oversight through DACs is grounded in the nature of data sharing, which allows downstream data uses that are not known at the time of the initial sample and data collection. Therefore, research ethics committees cannot foresee all downstream data uses when they approve the research protocol in the beginning. In that sense DACs are considered as an extra layer of oversight next to research ethics committees, which review the proposals in the beginning of the studies. In particular, DACs are established to receive data access requests from actual users and assess them for the purpose of approving or disapproving their access to data.[27] DACs are not mentioned in the GDPR, but their role in governance of data access is important. This can indeed be considered as part of research self-regulation in order to ensure data sharing and use is in line with the overarching principles and the relevant regulations.

[24] Joly et al. (2016), pp. 1150–1154.

[25] Philippakis et al. (2015), pp. 915–921.

[26] Wallace et al. (2014), pp. 149–157. See also: Ardeshirdavani et al. (2014).

[27] Lowrance (2012), p. 23.

DACs, function in different ways. As Lowrance illustrates, 'some of these groups are formally constituted, have terms of reference and hold regular meetings. Others, are casual, rarely meeting but existing to be consulted from time to time by the custodian and in a position to address serious problems should any arise'.[28]

The composition of DACs varies across the institutions. Ideally, such committees should be consisting of internal and independent members with expertise in technical, ethical and legal aspects of processing health and genomic data. Some have suggested establishing two-layer committee is beneficial, namely an advisory committees together with operational access committees. The advisory committees will be tasked with auditing the performance of the operational access committees, while the operational committee will be responsible for reviewing the access requests.

Moreover, the oversight committees, such as DACs and Research Ethics Committees, should be given the opportunity to assess the data access rules on a regular basis, and propose revision of the provisions when needed. This could ultimately strengthening effective operation of the organizational measures under Article 89(1). In addition, transparency of the data access governance could be considerably enhanced if adequate information dissemination policies are adopted. It is expected that the oversight bodies within the institutions provide information about the access review procedure, incoming data access requests and approved and disapproved requests to enhance transparency and facilitate external scrutiny. Furthermore, data access governance models should adopt mechanisms that hold users accountable.

5.2 Data Protection Impact Assessment and Appointment of Data Protection Officers (DPOs)

The GDPR sets further requirements in terms of governance of data processing when higher risks for the freedoms and the rights of the data subjects are perceived. One of the relevant organizational measures foreseen by the GDPR is to appoint a data protection officer (DPO) and conduct data protection impact assessment when specific conditions are met. The biobanks as entities that process health and genomic data should adhere to these provisions.

The Regulation in Article 37 provides a set of rules for designating the DPO when the processing of personal data within institutions meets certain criteria. According to the explanation provided by the European Data Protection Supervisor: 'the main task of the data protection officer is to ensure, in an independent manner, the internal application of the provisions of the Regulation in his/her institution. The data protection officer is also required to keep a register of all of the processing operations involving personal data carried out by the institution. The Register,

[28] Lowrance (2012), p. 23.

which must contain information explaining the purpose and conditions of the processing operations, should be accessible to any interested person.'[29] The appointment of a DPO must of course be based on her personal and professional qualities, but particular attention must be paid to her expert knowledge of data protection.

In addition, according to Article 35, a privacy impact assessment is necessary: 'where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purpose of the processing is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data.' Therefore, a broad scope for these data protection impact assessments is expected, which goes beyond compliance with the Regulation and privacy rights and includes consideration of a plethora of individual's fundamental rights. This will therefore provide an opportunity to take into account a broader range of concerns relating to the rights of individuals in processing personal data and not only those that are related to storage and safety. Article 35(b) adds that the data protection impact assessment shall in particular be required in cases where there is 'processing on a large scale of special categories of data referred to in Article 9(1)'.

The controller shall receive a data protection officer's advice (if he/she has been appointed) when carrying out a data protection impact assessment. Consequently, the data controller shall consult the supervisory authority prior to processing 'where a data protection impact assessment under Article 35 indicates that the processing would result in high risk.'[30] Article 35(9) also requires the data controller, where appropriate, to 'seek views of data subjects or their representatives on the intended processing'.

The impact assessment will therefore replace the previous obligation to notify the data protection authority, which was outlined by the Directive 95/46/EC on the protection of individuals with regard to the processing of personal data. This change was welcomed by commentators, who argued against the effectiveness of the previous notification requirement. As Townend argues: 'Although data controllers are required to register their activity with the relevant supervisory authorities and that authority has power to investigate and prosecute breaches of the data subject's rights, the sheer amount of processing that goes on within any jurisdiction at any given time makes it impossible for a supervisory authority to be seen as the primary protector in the system'.[31] In turn, the new requirements will see a shift towards the accountability of the controllers and reinforce their role in establishing adequate safeguards in the course of the processing, not only limiting it to the outset of the project.[32] This could also draw the attention of the data controllers towards the

[29] European Data Protection Supervisor. https://edps.europa.eu/data-protection/eu-institutions-dpo_en.

[30] Article 36, GDPR.

[31] Townend (2016), pp. 128–142.

[32] de Hert and Papakonstantinou (2016), pp. 179–194.

ethical concerns associated with data processing and take the concerns into account in the design of the processing.

5.3 Data Access Agreements and Data/Material Transfer Agreements

Contractual agreements are essential operational instruments intended to legally bind the parties to specific rules ensuring adequate individuals' privacy protection throughout personal data processing. Such contracts can take many forms and be labelled differently such as 'Data Access Agreements' (DAAs), 'Data Transfer Agreements' (DTAs) or 'Confidential Data Agreements' (CDAs). Personal data protection measures are also included within other special research agreements, in particular within the so-called 'material transfer agreements' (MTAs) when biological samples are also transferred. In particular, it is widely recommended to include in MTAs provisions regarding samples' quality, transportation, conditions and restrictions of use (e.g. derivations of original material) and storage (biosafety/biosecurity).

The nature and scope of the contract can vary depending on the internal practices of operators or applicable national legal framework, the requester's processing operation and purposes, the database governance model (cf. supra) and on the cross-border features of the access (intra-EU or including outside-EU elements). For example, where data is managed within a closed controlled system (e.g. digital data analysis platform), an access agreement could take the form of terms and conditions in the view of the applicable regulations. In addition, a decentralized infrastructure could rely on a general Access policy having a contractual value. For example, BBMRI-ERIC[33] provides such template while allowing its members to adopt specific and compliant contractual activities to frame collaborations.[34]

The legal qualification of the parties to such agreements is context-dependent and needs a case-by-case analysis of the role and activities of each stakeholder. Access could be requested in a framework of a research collaboration with the biobank or by an external researcher to conduct an independent research project. Thereby, the contract will define a controller-processor relationship or a joint-controllers relationship. This is in line with the GDPR that requires setting up a contract for organizing joint-controllers[35] and/or controller-processor relationships[36] in terms of duties and rights in processing data.

[33] BBMRI-ERIC (2018).

[34] B3Africa, Checklist: For a good governance of transcontinental collaborative biobank research. http://biobanklearning.iarc.fr/course/checklist-elsi/#llms-lesson-locked. Accessed 9 May 2019.

[35] Article 26, GDPR.

[36] Article 28, GDPR.

The data access agreements usually include negotiable and non-negotiable provisions. Contracts shall echo and respect the will of the initial donor and facilitate the exercise of the donors' rights. The parties shall commit to respect confidentiality and plan cooperation procedures, in particular regarding personal data breach notifications. The agreement must also clearly describe any restriction specified by the initial controller during the deposit of the data/sample in the biobank or imposed by the biobank policy based on a legitimate interest (e.g. regarding onward transfers possibilities, the return of the data/samples or destruction, intellectual property issues). For ensuring proper legal security, agreements must include information about the applicable laws and dispute resolution mechanisms, including out-of-court proceedings. Financial provisions could also be included but should not be indexed on the intrinsic personal data or sample value but on the necessary investments performed for ensuring samples or data quality, integrity and FAIRness for example.

In addition, the GDPR is setting specific conditions when transferring data/and samples to non-EU countries. Accordingly, materials can only be transferred to a third entity in a country that ensures an appropriate level of protection of individuals' rights and freedoms compared to the one guaranteed within the EU. Therefore, such a transfer can be permitted where it is based on an adequacy decision adopted by the European Commission after analysis of a country general and sectorial legislation,[37] or where appropriate safeguards are in place.[38] This is including the use of binding corporate rules (applying to cross-border personal data transfers in a group of undertaking or a between entities of a multinational enterprise), or of standard contractual clauses adopted by the European Commission[39] (provided that they are not modified, otherwise the competent supervisory authority should be consulted to validate the adapted clauses), the respect of an approved Code of Conduct or the use of an approved certification mechanism together with binding and enforceable commitments of the controller or processor in the third country to apply the appropriate safeguards, including as regards data subjects' rights.

In exceptional circumstances, in the absence of an adequacy decision and of appropriate safeguards a transfer shall take place only if one of the conditions of Article 49 GDPR are met. This includes situations where the data subject has explicitly consented to the transfer, after having been informed of the possible risks of such transfers for the data subject due to the absence of an adequacy decision and appropriate safeguards, or the transfer is necessary for protecting the vital interests of the data subject, or is necessary for important reasons of public interest recognized in the Union or relevant Member State law (e.g. fight against cross-border public health threats), or is made from a public register intended to provide

[37] Article 45, GDPR.

[38] Article 46, GDPR.

[39] European Commission. Model contracts for the transfer of personal data to third countries. https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/model-contracts-transfer-personal-data-third-countries_en. Accessed 9 May 2019.

information to the public and which is open to consultation either by the public in general or by any person who can demonstrate a legitimate interest.

6 Conclusions

In developing data access rules and governance models biobanks could be assisted by soft law measures, which have traditionally had considerable importance in the field. It seems that GDPR leaves considerable room to operationalize its provisions through these soft law measures. One area that soft law measure can be useful is in elaborating on what organizational measures should be when processing data under research exemptions. In particular, such measures can provide guidance on the adequate models of data governance, oversight bodies, data access rules and implementation of data protection best practices.

Oversight bodies can be considered as a crucial part under organizational measures. In particular, oversight bodies such as ethics committees and data access committees are in the good place to hold control over the access and use of data. It is important to ensure the existing and emerging oversight bodies are equipped with adequate expertise regarding using and sharing genomic data and are aware of the associated informational risks. In order to achieve this, soliciting the attitudes of the involved parties regarding the associated risks would be necessary. Thereby, the overall governance of personal data processing will go beyond legal requirements and will take into account the pertinent individual or social concerns that may not be explicitly outlined in the legal provisions. That said, DACs often lack adequate tools to keep ongoing oversight on actual use of data once data access has been granted. Such limitations on the oversight on data access should be taken into considerations, when assessing the potential risks and the adequacy of the current oversight tools and mechanisms.

Moreover, the oversight of personal data processing by competent authorities should keep pace with recent developments in the field of data science, bioinformatics and genetics, among others. The risks associated with emerging technologies and the safeguards in protecting the privacy of data subjects should be treated as moving targets. Otherwise, the safeguards will become obsolete and unable to safeguard data subjects in an adequate fashion.

Finally, increasing cross-border data sharing underlines the importance of the harmonization of legal frameworks concerning personal data protection. One of the main goals of the Regulation has been to achieve this by harmonizing the personal data protection landscape across EU. However, concerns remain regarding the real impact of the Regulation on unifying the national regulations towards processing health and genetic data for research purposes, across Member States. Arguably, the Regulation still leaves room for varying interpretations, for instance, concerning the safeguards that should be established and also in setting further conditions for processing data on the basis of the research exemption provisions. This may challenge development of European sample repositories and data sharing platforms, as

different safeguards may be required to be adopted for samples/data collected in different member states.

Acknowledgements L.M has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 753531.

References

Ardeshirdavani A, Souche E, Dehaspe L et al (2014) NGS-logistics: federated analysis of NGS sequence variants across multiple locations. Genome Med 6:71

Article 29 Data Protection Working Party (2014) Opinion 05/2014 on anonymization techniques. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm. Accessed 27 May 2019

Article 29 Data Protection Working Party (2018) Guidelines on consent under regulation 2016/679 (wp259rev.01). https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051. Accessed 29 May 2019

B3Africa, Checklist: For a good governance of transcontinental collaborative biobank research. http://biobanklearning.iarc.fr/course/checklist-elsi/#llms-lesson-locked. Accessed 27 May 2019

BBMRI-ERIC (2018) BBMRI-ERIC policy for access to and sharing of biological samples and data. Available on line at: http://www.bbmri-eric.eu/wp-content/uploads/AoM_10_8_Access-Policy_FINAL.pdf. Accessed 27 May 2019

Cook-Deegan R, McGuire AL (2017) Moving beyond Bermuda: sharing data to build a medical information commons. Genome Res 27:897–901

Court of Justice of the European Union (CJEU), Case C-582/14: Patrick Breyer v Bundesrepublik Deutschland. http://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=EN. Accessed 27 May 2019

de Hert P, Papakonstantinou V (2016) The new general data protection regulation: still a sound system for the protection of individuals? Comp Law Secur Rev 32(2):179–194

Dyke SO, Kirby E, Shabani M et al (2016) Registered access: a 'Triple-A'approach. Eur J Hum Genet 24:1676–1680

European Commission. Model contracts for the transfer of personal data to third countries. Available online at: https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/model-contracts-transfer-personal-data-third-countries_en. Accessed 27 May 2019

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

Joly Y, Dyke SO, Knoppers BM et al (2016) Are data sharing and privacy protection mutually exclusive? Cell 167:1150–1154

LERU (2016) Policy Brief: the new EU general data protection regulation: why it worries universities and researchers. Available at: https://www.leru.org/files/The-New-EU-General-Data-Protection-Regulation-Why-It-Worries-Universities-and-Researchers-Full-paper.pdf. Accessed 27 May 2019

Lowrance WW (2012) Privacy, confidentiality, and health research. Cambridge University Press, Cambridge

Marelli L, Testa G (2018) Scrutinizing the EU general data protection regulation. Science 360(6388):496–498

Mascalzoni D, Bentzen HB, Budin-Ljøsne I et al (2019) Are requirements to deposit data in research repositories compatible with the European Union's general data protection regulation? Ann Intern Med 170(5):332–334

Nuffield Council on Bioethics (2014) Collection, linking and use of data in biomedical research and health care

Paltoo DN, Rodriguez LL, Feolo M et al (2014) Data use under the NIH GWAS data sharing policy and future directions. Nat Genet 46:934

Philippakis AA, Azzariti DR, Beltran S et al (2015) The matchmaker exchange: a platform for rare disease gene discovery. Hum Mutat 36:915–921

Rodriguez LL, Brooks LD, Greenberg JH et al (2013) The complexities of genomic identifiability. Science 339:275–276

Shabani M, Borry P (2016) "You want the right amount of oversight": interviews with data access committee members and experts on genomic data access. Genet Med 18:892–897

Shabani M, Borry P (2017) Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. Eur J Hum Genet 26:149–156

Shabani M, Marelli L (2019) Re-identifiability of genomic data and the GDPR: assessing the re-identifiability of genomic data in light of the EU General Data Protection Regulation. EMBO Rep:e48316

Shabani M, Dyke SO, Marelli L et al (2018) Variant data sharing by clinical laboratories through public databases: consent, privacy and further contact for research policies. Genet Med 21:1031–1037

Townend D (2016) EU laws on privacy in genomic databases and biobanking. J Law Med Ethics 44:128–142

U.S. Department of Health & Human Services (2012) Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) privacy rule. Available at: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html#standard. Accessed 27 May 2019

Wallace SE, Gaye A, Shoush O et al (2014) Protecting personal data in epidemiological research: DataSHIELD and UK law. Public Health Genomics 17:149–157

Woolley JP, Kirby E, Leslie J et al (2018) Responsible sharing of biomedical data and biospecimens via the "Automatable Discovery and Access Matrix"(ADA-M). NPJ Genomic Med 3:17