

The European General Data Protection Regulation: challenges and considerations for iPSC researchers and biobanks

Michael Morrison¹, Jessica Bell^{1,2}, Carol George³, Shawn Harmon³, Megan Munsie⁴ & Jane Kaye^{*,1,2}

¹Nuffield Department of Population Health, HeLEX – Center for Health, Law & Emerging Technologies, University of Oxford, Ewert House, Ewert Place, Banbury Road, Oxford OX2 7DD, UK

²Melbourne Law School, University of Melbourne, Parkville, VIC 3010, Australia

³School of Law, Kenyon Mason Institute for Medicine, Life Sciences & Law, University of Edinburgh, Old College, South Bridge, Edinburgh EH8 9YL, UK

⁴Department of Anatomy & Neuroscience, Stem Cells Center for Stem Cell Systems, University of Melbourne, Parkville, VIC 3010, Australia

* Author for correspondence: jane.kaye@law.ox.ac.uk

Increasingly, human induced pluripotent stem cells (iPSC) and their associated genetic and clinical information are being used in a wide range of applications, with large biobanks being established to support and increase their scientific use. The new European General Data Protection Regulations, which comes into effect in 2018, will have implications for biobanks that generate, store and allow research access to iPSC. This paper describes some of the challenges that iPSC biobanks face and suggests some points for the development of appropriate governance structures to address these new requirements. These suggestions also have implications for iPSC research in general.

First draft submitted: 3 May 2017; Accepted for publication: 26 July 2017; Published online: 4 October 2017

Keywords: biobanks • data protection • governance • induced pluripotent stem cells • stem cells

It was only a decade ago that the first report of cell reprogramming to produce an ‘embryonic cell-like state’ using cells from adults, rather than the more socially controversial embryos, was reported [1]. Since then, the technology to create these induced pluripotent stem cells (iPSC) has become more widely used, and the technique has been incorporated into a range of major translational applications [2]. Human iPSC (hiPSC) are currently being investigated as a source of cells for regenerative medicine [3], as tools for *in vitro* modelling of human development and disease, and for cell-based assays of small molecule drug candidates [4–6]. In these latter applications, iPSC are viewed as especially valuable because they can be derived from adults or children who have particular conditions, enabling disease-specific cell lines to be created, or from healthy volunteers to generate well-defined control lines. A number of high-profile endeavors – the European Bank for induced pluripotent stem cells (EBiSC), the Stem Cells for Biological Assays of Novel drugs and predictive toxicology consortium (StemBANCC), the New York Stem Cell Foundation stem cell bank and the Human Induced Pluripotent Stem Cell initiative (HipSci) – are working to produce, bank and disseminate human iPSCs to support further investigation of these translational applications [7–10].

Unlike human embryonic stem cells, human iPSC usually originate from individuals from whom additional phenotypic, clinical and behavioral data (such as family history, age of disease onset, medications, diagnostic results etc. [11]) may be accessed. Unless the iPSC are derived from fetal tissue or a deceased donor, these individuals are ‘natural or legal persons’ and the data are ‘personal’ and ‘health’ data, which are considered especially sensitive and deserving of legal protection. All personal data are regulated by data protection law, meaning donors are entitled to a high degree of privacy protection and to security of the data associated with the cell line.

European data protection law is set to change. In May 2018, the current European Data Protection Directive (Directive 95/46/EC) [12] will be replaced with the new General Data Protection Regulation (GDPR; or Regu-

Box 1. Main changes resulting from the replacement of Directive 95/46/EC with the General Data Protection Regulation

- Higher fines for firms contravening the new Regulations based on global turnover
- New requirements for organizations to appoint a Data Protection Officer if the organization handles ‘significant’ amounts of sensitive data
- Genetic data are now explicitly recognized as ‘sensitive personal data’
- ‘Privacy by design’ approach to data protection strongly advocated, with data protection safeguards to be built into organizations’ products and services from the earliest stage of development and ‘privacy-friendly’ techniques such as pseudonymization, encryption and anonymization promoted
- New direct obligations on data processors, including maintaining certain records of all processing activities and taking appropriate security measures to protect personal data
- Expanded territorial reach as the General Data Protection Regulation places obligations on non-EU organizations who process or monitor data of EU citizens
- Broad consent for scientific research is specifically allowed although more specific consent options should be offered if compatible with the research
- Data being used for research are exempted from some of the new data subject rights such as the right to be forgotten and the right to object to processing of ones’ data, but must be collected and stored subject to appropriate safeguards enforcing the principle of data minimization
- Whereas some national interpretations of the Directive allowed data being used for research to be stored indefinitely, the Regulation stipulates research data should be stored ‘for as long as necessary’

lation) [13]. The scope of the GDPR includes ‘pseudonymized data’; where the sample donor is de-identified, but a key that allows a connection to the original tissue donor to be made is maintained. Data protection changes will apply whether iPSC are derived from ‘new’ donated material or existing tissue samples that are not fully anonymous. Projects based in European countries, including EBiSC, StemBANCC and HipSci, are required by European laws on the traceability of biological material to pseudonymize rather than anonymize data and samples so their operations automatically come under the scope of the new Regulation. An example traceability system implementing pseudonymization for hiPSC biobanking is described in [14]. All of this means that existing arrangements in place for sharing human embryonic stem cell lines and associated data, including SNP genotyping [15], are not automatically adequate and appropriate for disseminating well-characterized iPSC lines.

Given the imminence of the GDPR – which strengthens a regime already considered one of the most stringent in the world [16] – it is timely and appropriate to review the data protection implications of making iPSC and associated genetic and clinical data available to the global scientific community. The aim of this paper is to offer some preliminary views on the impact of changing data protection law for iPSC researchers and iPSC-based biobanks. First, we identify some of the core features of the GDPR. Second, we identify key matters for consideration when designing governance arrangements for iPSC biobanks. Third, drawing on the standards articulated in the GDPR, we offer some recommendations, based on our collective experience in the biobank setting, as to how to proceed.

The General Data Protection Regulation

The GDPR updates existing European data protection law, replacing EU Directive 95/46/EC on data protection. The major differences between the Regulation and existing European data protection law are summarized in Box 1.

Under the terms of the Directive, European Member States were entitled to implement data protection requirements into national law in ways that allowed for differences between the jurisdictions [17]. The Regulation, by contrast, should be directly transposed into the national laws of EU Member States, in principle producing harmonization of European data protection legislation. However, the GDPR does allow for some national variations (derogations) in certain circumstances [18] so it is likely that some differences between Member States will persist after 2018. For example, each state can set a minimum age for consent to use of data at between 13 and 16 years of age. Prior to implementation the extent and impact of national variations cannot be known, but this is an area that will require further monitoring. The UK government has confirmed, through its Information Commissioner’s Office, that it will also implement the GDPR despite the decision to exit from the EU [18]. Therefore, the GDPR will apply uniformly to all researchers, biobanks and institutions using iPSC cells located within Europe. Key aims of the GDPR, articulated in Recitals 7, 9, 10 and 11, and Article 1, are described in Box 2.

The pursuit of these aims is guided by the fair processing principles set out in Article 2 of the GDPR and described in Box 3.

Box 2. Key aims of the General Data Protection Regulation

- Ensure a consistent and high level of protection of natural persons while also removing the obstacles to flows of personal data within the EU
- Harmonize the level of protection of the rights and freedoms of natural persons with regard to the processing of such data across the EU
- Clarify the obligations of data processors and ensure strong enforcement

Box 3. Fair processing principles set out in the General Data Protection Regulation

- The purposes for processing must be legitimate, specified and limited
- The processing practices must be lawful, fair and transparent
- They must involve appropriate security to safeguard privacy and data integrity
- The data must be adequate, relevant and limited (data minimization), and must be accurate and maintained as identifiable only for as long as is necessary given the purposes of collection

In all cases, data controllers have primary responsibility, and will be held accountable for ensuring compliance. Given the tenor of the GDPR, we can assume that data controllers (here universities, hospitals and companies operating a biobank or processing large amounts of personal data for other research purposes) will be expected to meet fairly strict standards when it comes to receiving, holding and distributing (processing) cell lines and associated data.

However, EU law is guided by the principle of proportionality [20], which allows the application of the law to be adapted to particular circumstances, in this instance biobanking. Article 24 of the Regulation states that, taking into account the nature, scope, context and purposes of processing, as well as the likelihood of breaches or failures, and the severity of harms resulting therefrom, controllers shall implement appropriate technical and organizational measures. Article 25 acknowledges context and cost implications of decision-making when designing data protection measures.

Biobanks may also be exempt from a number of provisions of the GDPR, in particular under exemptions provided to permit the processing of personal data for purposes of scientific research. Such provisions allow personal data to be stored for longer periods than would otherwise be permitted, provided that: they will be processed solely for scientific research purposes in accordance with GDPR Article 89(1), and; that the data will be subject to implementation of technical and organizational measures required by the GDPR. The Regulation also retains a presumption of ‘compatibility of use for research purposes’, which allows personal data collected for purposes other than scientific research to be repurposed for research as long as there is a valid legal ground for the original processing in EU or Member State law. Finally, universities may be exempt from the requirement to protect some of the rights of data subjects accorded in the Regulation, where to exercise these rights would prevent or seriously impair the possibility of the research being carried out.

A number of provisions of the GDPR apply to anyone who generates human-derived iPSC and associated data, and/or wishes to bank and distribute them. The Regulation also applies to anyone who wishes to work with data derived from EU citizens, whether or not the data user (processor) is itself based in the EU. The GDPR is therefore relevant for anyone using human data, including scientists and biobank managers, in both EU and non-EU jurisdictions. For some of the key terms, as they are defined in the GDPR, and their relevance to iPSC banking, see Table 1.

The General Data Protection Regulation & key operational matters

As compliance with data protection law is a function of biobank governance, it implicates a number of matters, which are simultaneously operational, legal and in some cases ethical in nature. Here we highlight several such matters for which the GDPR has particular relevance.

Consent

Consent is the most common mechanism by which we enable individuals to exercise their autonomy in relation to medical research, though its limits in terms of practice must be acknowledged. The GDPR requires a legal basis for processing personal data and consent is one such legal basis. Consent must be obtained before personal

Table 1. Key terms in the General Data Protection Regulation and their relevance to induced pluripotent stem cell banking.

Term	Meaning in the GDPR	Relevance to iPSC Research
Data subject	An identifiable natural person who can be identified, directly or indirectly, especially 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	European regulations on the traceability of biological material [21,22], mean that, for quality and safety purposes, all iPSC derived from new samples must be traceable to the sample donor. This means that data associated with iPSC and derived from them can be pseudonymized but must not be fully anonymized. iPSC derived from new or existing samples where the sample donor is not fully anonymized therefore relate to identifiable legal/natural persons and so fall within the scope of the GDPR. This requirement is mirrored by WHO Guidelines on traceability of human organs, tissues and cells [23] meaning pseudonymization is also recommended practice outside the EU
Data processing	Any operations that are performed on personal data or sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction	Collection, storage, editing, analyzing or otherwise working with genetic or patient data associated with an iPSC line counts as data processing
Data processor	A natural or legal person, public authority, agency or other body that processes personal data	Any researcher or entity that collects, stores, edits, analyses or otherwise works with genetic or patient data associated with an iPSC line counts as a data processor
Data controller	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	Any institution, such as a university, that runs a biobank or employs scientists who work with personal data (including pseudonymized data associated with iPSC or other cell lines) counts as a data controller
Supervisory authority	An independent public authority in each EU Member State tasked with the responsibility of monitoring and enforcing the correct application of the Regulation. All countries signed up to the GDPR must nominate a supervisory authority. Some states may choose to split different areas of this responsibility between two or more public bodies	Data controllers and data processors who work with iPSC are ultimately responsible to the supervisory authority for their handling of sensitive personal information associated with, or derived from, the cell lines. Supervisory authorities are also involved in monitoring cross-border flows of data and can require that data not be provided to a specified non-EU country
Data protection officer	An individual appointed by an institution involved in processing personal data on a large scale to oversee compliance with data protection law.	A data protection officer would be able to help biobanks, as well as academic or commercial scientists working with iPSC, to meet their data protection obligations
Personal data	Any information relating to an identified or identifiable natural person (data subject)	Fully anonymized data do not count as personal data. Pseudonymized data are re-identifiable and so count as personal data caught by the GDPR
Data concerning health	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	All three categories of data are considered especially sensitive, warranting a higher level of protection than other types of information. Data concerning health and/or biometric data may accompany iPSC derived from patients or healthy volunteers. Some elements of genetic data are also likely to be provided with the iPSC line to characterize it, or will be derived from the biological sample by recipients of the cell line. Once genetic information is extracted from the sample, even if it is not whole genome/exome sequence data, it becomes subject to data protection law
Genetic data	Personal data relating to the inherited or acquired genetic characteristics of a natural person which give unique information about the physiology or health of that natural person and which result, in particular, from an analysis of a biological sample from the natural person	–
Biometric data	Personal data resulting from specific technical processing relating to the physical, physiological or behavioral characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or fingerprint data	–

GDPR: General Data Protection Regulation; iPSC: Induced pluripotent stem cell.

Box 4. Updated record keeping requirements for data controllers.

- Name and contact details of the controller;
- Purposes of the processing;
- Description of categories of data subjects and personal data;
- Categories of recipients to whom personal data have been, or will be, disclosed, including in third countries or international organizations;
- Transfer of personal data to a third country or international organizations;
- Where possible, envisaged time limits for erasure of the different categories of data; and
- Where possible, a general description of the technical and organizational security measures referred to in Article 32(1).

information about an individual (a data subject) can be collected and processed (Article 7). It is mandated that consent should be informed and explicit. This applies to consent provided by an individual participant, or by a proxy on behalf of someone who is not capable of giving valid consent (e.g., for collection of samples from pediatric cohorts or patients with dementia). The provisions for research make it clear that the use of a broad consent, where participants agree that their samples and data can be made available, subject to a suitable governance framework being in place, is permissible in biobanking research. The requirement to be explicit about what is being consented to, however, has implications for how consent is presented and collected in projects or biobanks recruiting iPSC donors. At minimum, consent forms should contain separate statements (to be individually signed/initialed by the participant) reflecting agreement that a biological sample can be collected, can be reprogrammed, can be stored and can be shared together with the personal and/or medical data of the participant. When proposing to derive human iPSC from tissue samples that were obtained previously in support of a research plan that did not include iPSC derivation, then the validity of the existing consent for this new purpose will need to be reviewed by an appropriate body such as a biobank governance board or a research ethics committee [24].

Genetic and biometric data can be considered ‘inherently identifying’ because they relate to a particular individual. The GDPR imposes a general prohibition on processing of data wherever identification of an individual data subject is possible, unless explicit consent has been given (Article 9). It is therefore advisable to seek explicit consent for the collection and processing of genetic or biometric data, and researchers should avoid giving tissue donors a guarantee of absolute anonymity or privacy. To promise that the donor cannot and will never be identified is to encourage an unrealistic expectation, and could expose researchers to liability if a data subject is identified (e.g., through a data breach) [11,25]. Academic studies that generate iPSC should also consider the consent requirements associated with making these cell lines available to the wider research community. It is important to be clear about how the data will be used and distributed, so recruiters should ensure that participants give permission to share cell lines and data with researchers in other countries, the private sector and consent to post-study deposition of the lines and data in a biobank such as EBISC.

Record keeping & accountability

The principles of accountability and transparency have been strengthened by the GDPR and are now central to data protection. These principles have increased the administrative responsibilities of data processors and controllers, who must now ensure that they have effective governance mechanisms in place. They must be able to demonstrate compliance both internally, and externally, to the relevant supervisory authority, which, in the case of England and Wales, is the Information Commissioner’s Office. Where more than one EU country is involved in a project, there will be an agreed lead supervisory authority. Importantly, the GDPR imposes, for the first time, direct responsibilities and penalties on data processors, in addition to data controllers. Under Articles 28 and 30, processors are required to maintain records of processing activities, and to take appropriate security measures to protect personal data ‘to ensure a level of security appropriate to the risk.’

Under the rubric of transparency and accountability, Article 30 stipulates that data controllers must keep the records listed in Box 4.

Biobanks and researchers using standard information management tools, such as a Laboratory Information Management System, commonly maintain records of the location of all biobank samples/stock and the period of time they are expected to be held there. The GDPR – particularly Article 30(d) and (e) – suggests that further records, concerning origin of data associated with each cell line (source, collection time, institutional or individual

collector) may also need to be kept. These standards would apply to online catalogues and systems like the Human Pluripotent Stem Cell Registry [26] and the Stem Cell Database [27].

Another GDPR requirement is that controllers and processors must keep records up to date (Article 30). In addition to being an explicit rule, this requirement for stable and durable linkages is a sensible practice relevant to the longevity of the iPSC as a communal resource. Under this rule, biobanks and iPSC researchers who generate cell lines and collect clinical data must find ways to ensure a long-term, stable means of access to the pseudonymization key that links individual donors and their samples/data. If this is left to individual scientists, clinicians, managers or hospital staff who have had personal contact with the biobank, for example, as material or data depositors, these connections are likely to be lost as projects end, contact details change and people move on over time.

One way to achieve this stability and durability is suggested by Article 37(1) of the GDPR, which obliges institutions processing sensitive and personal data on a large-scale to appoint a Data Protection Officer (DPO), whose task is to assist the monitoring of internal compliance with the GDPR. Given the nature of biobanks such as EBiSC and New York Stem Cell Foundation, and the large numbers of iPSC intended for production and dissemination by them, the appointment of a DPO is a reasonable and pragmatic undertaking [28]. For biobanks and research groups who produce and share human iPSC, part of the role of the DPO could be to serve as custodian for the pseudonymization keys, and to act as a contact point for traceability and re-contact requests. Although organizations with less than 250 employees are exempt from the obligation to appoint a DPO, it should be noted that from the perspective of a biobank, the organization in question may be a larger entity within which the bank is situated; in other words, if the biobank is housed in or associated with a hospital or university, the number of employees for regulatory purposes will be that of the broader hospital or university [29].

Data minimization

The principle of data minimization is strongly embedded in the GDPR. Article 5(1)(c) states that personal data shall be ‘adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (data minimization).’ The importance of this principle to the new regime is indicated by its reinforcement in both Article 25, which addresses data protection by design, and Article 89(1), which addresses safeguards and derogations where data are used specifically for research.

The practical achievement [30], and indeed the propriety of the pursuit, of data minimization present challenges in the biobank setting. The notion of a quantitative restriction that prohibits collection of more data than are strictly necessary for the task at hand is a concept inherently problematic for all scientific research, but particularly for undertakings that aim to generate resources for presently unanticipated future investigations. Increasingly, funding bodies and publishers are requiring researchers and research institutions to share their data as a criterion of receipt of funding or publication of results [31]. Further, the key aim of accessibility held by large iPSC banks, which would make iPSC cells and data ‘maximally available’ to as many scientists as possible, also seems at odds with this principle, in so far as it means keeping the amount of processing and the number of processors to a minimum.

At present, it is recommended that biobanks take steps to ensure that the data they collect are related to the purposes ‘for which they are processed,’ which will be primarily research (broadly conceived) and that this is appropriately communicated in the consent process. For example, samples and data from patients with neurodegenerative disease might reasonably include data on age of onset, as this is clinically relevant to the research topic even though this information will not necessarily be utilized in every research project involving the samples. Scientists and biobanks should also de-identify the data to the extent that this does not interfere with research objectives. It is also advisable that those biobanks that collect samples and associated data for long-term and uncertain future uses should review existing data protection provisions for relevance each time they plan a further data collection initiative, or when a significant amount of new data are added. Ethical oversight bodies, such as data access committees, research ethics committees and institutional review boards, should also be vigilant in considering whether particular requests for data access – for example, to provide ‘all data’ on a particular data subject – warrant additional safeguards to ensure that the anonymity of the subject is secure.

Transfer of samples/data outside the EU

International transfers of data to third countries or international organizations are a matter of significant uncertainty for biobanks intending to promote global access to their resources, and they hold an ambiguous position within the GDPR. Despite the Recital 101 affirmation that international flow of personal data is necessary for the expansion of trade and cooperation, Article 44 of the GDPR begins with a prohibition of such international transfers of

personal data, on grounds that such movement of personal data may jeopardize both the ability of natural persons to protect themselves from unlawful use or disclosure of that information, and the ability of EU authorities to pursue complaints or conduct investigations relating to activities outside their borders.

The primary exception to the prohibition, articulated in Article 45, is a Commission decision, with effect for the entire EU, that there is an ‘adequate level of protection’ of the fundamental rights of the data subject in the foreign jurisdiction (i.e., an opinion that the foreign protections are essentially equivalent to that ensured within the EU). Although an EU registry of data-friendly jurisdictions might ultimately increase the efficiency of iPSC dissemination, and promote harmonization of global data protection, it is unlikely to provide an immediate avenue for validation of dissemination of cells and data by international iPSC biobanks. Commission decisions will take time, as they are made by means of an implementing act following an assessment that takes into account the legal environment, the existence and functioning of independent supervisory authorities and international commitments of the foreign jurisdiction, particularly in relation to protection of personal data.

The more pertinent question for biobanks may be whether standard Material Transfer Agreements, Codes of Conduct or other mechanisms contemplated by the GDPR are capable of providing ‘appropriate safeguards’ for the legitimation of routine international transfer of data (Article 46). Some basis for proceeding is the existing ‘model contract’ clauses for transfer of personal data outside the EU that were produced to help data controllers comply with the previous EU data protection Directive [32,33]. In the absence of an adequacy decision by the Commission under Article 45.3, a controller or processor may transfer personal data to a third country or an international organization only if; the controller or processor has provided ‘appropriate safeguards’; and enforceable rights and effective legal remedies are available to data subjects.

Article 46(2) offers a list of ‘appropriate safeguards’ – including binding instruments between public authorities, binding corporate rules, standard data protection clauses adopted by the Commission or a supervisory authority, an approved code of conduct or certification mechanism – but each of these require the prior approval of the Commission or a relevant independent supervisory authority established under Chapter VI. Further, there is no instruction offered with respect to how to ascertain the availability of enforceable rights and effective remedies for data subjects in the foreign jurisdiction, or who is responsible for doing so.

Conclusion

International circulation of biologicals and associated data is considered essential for the advancement of basic research and clinical translation of stem cell science. Biobanks and registries for disseminating well-characterized iPSC are an important part of this endeavor, but they also raise particular issues for data protection, as most iPSC are derived from individuals with data protection rights. The forthcoming European GDPR represents a change to the legal requirements for data protection. It is therefore important to communicate how these changes affect the responsibilities of biobanks, and researchers accessing biobanks, so that they can remain in compliance with the law during the course of their work. This is doubly relevant because the GDPR extends the legal reach of European law to any person or institution processing data derived from EU citizens, regardless as to where that person or institution is located [16].

Some of these responsibilities can be dealt with relatively easily by putting in place data protection policies for projects, institutions and consortia. Such policies might ensure, for example, that no sensitive data leaves a secure site unless it is stored on an encrypted device, or that contracts and legal agreements contain provisions requiring anyone with access to pseudonymized personal data to refrain from attempting to identify any data subject. Other elements of the GDPR look likely to increase the bureaucratic burden of sharing, at least on a large scale, human-derived iPSC cells and data and will require additional administrative support such as the appointment of specialist Data Protection Officers.

The most challenging of the GDPR provisions are those in which the aims of the Regulation seem least aligned with those of biobanking: data minimization and transfer of samples and data outside the EU. Data minimization might be mitigated by a move toward enabling greater, even ongoing, contact with sample donors, allowing researchers to go back and collect additional data if, as, and when it is needed instead of treating data collection as a one-off event [34]. A move toward digital tools for consent and engagement may offset some of the administrative burden of sustained interaction [35,36]. The issue of transfer of data outside the EU is potentially the most challenging development. It should be noted that ‘transfer’ outside the EU includes digital transfer of data via ‘cloud computing’ services, if the data are stored on a server that is located in a non-EU country. The Regulation transforms into a formal legal responsibility that was previously an internal process: assessment by biobank staff of the legitimacy of

Box 5. Summary data protection recommendations for induced pluripotent stem cell biobanks

- Consent forms for recruiting donors of induced pluripotent stem cell (iPSC) originating tissue need to contain separate, explicit consent for donation of the biological sample for reprogramming and for the collection and processing of personal data and genetic data that will be used to characterize the cell line
- A broad consent can be collected, but biobanks and researchers should still consider adding explicit consent provisions to allow iPSC lines and associated data to be made available to the wider scientific community, including commercial companies
- Institutions need a long-term plan for maintaining connections between pseudonymized data and sample donors. This connection needs to be durable and enable access from subsequent data controllers, even if the samples/data are produced by a fixed-term project or consortium
- Data controllers and processors need to be aware of their increased administrative responsibilities in order to ensure compliance with the record keeping requirements of the General Data Protection Regulation accountability and transparency provisions
- Data controllers should consider implementing a 'privacy by design' approach to all operations in order to meet privacy, security and data minimization commitments
- In order to comply with the principle of data minimization, biobanks and scientists deriving human iPSC should ensure that collected data are related to the purpose for which they are processed, in other words, research, broadly conceived. This should be reviewed if a significant tranche of additional data are collected. This may require putting in place mechanisms to recontact donors to update their data or provide additional information
- Personal data on EU citizens, including pseudonymized data, may only be transferred to data processors in non-EU countries if these have 'adequate protection' for data security in place. In the long term, the Commission is expected to rule on the data protection adequacy of various non-EU states, but in the short term data processors and controllers are advised to make use of formal contracts such as Material Transfer Agreements that employ standard data protection clauses of the kind provided on the EU and Information Commissioner's Office websites

a person or institution who requests access to a cell line and associated data [37]. This change is problematic because the responsibility will be conferred before the necessary range of decision-making tools (in the form of a registry of EU approved or prohibited data destinations) are made available and in this regard the only option seems to be to stick to current best practice for data transfer while pushing the Commission to clarify its adequacy decisions on relevant countries and seeking additional resources to deal with this requirement.

In (Box 5), we offer a list of key recommendations to assist biobanks and researchers who work with human-derived iPSC to comply with changing data protection law.

Future perspective

Biobanks are important resources for translational research but they also face challenges to sustainability [38,39]. It is likely that over time there will be first increasing collaboration between different iPSC banks driving harmonization of standards, including those for data protection, and ultimately international networks of biobanks working to facilitate the international flow of cell lines and data across jurisdictions [40]. There is precedent in initiatives such as Promoting Harmonization of Epidemiological Biobanks in Europe and the Biomolecular Resources Research Infrastructure–European Research Infrastructure Consortium [41]. Pluripotent stem cell banks founded in or after 2018 will need to be designed to be GDPR compliant from the outset.

At the same time meeting quality, safety, data protection, traceability and other commitments takes considerable administration work and is likely to result in high running costs for hiPSC banks. To meet these costs cell banks may have to leverage the value of well-characterized iPSC to the pharmaceutical industry through longer term engagement between public and private sectors.

Failure to adequately address issues associated with data protection and related aspects of biobank governance run the risk of not only stifling research using iPSC but public trust and support for this important facet of biomedical research.

Disclosure

This publication reflects only the author's views and neither the Innovative Medicine's Initiative Joint Undertaking nor the European Federation of Pharmaceutical Industries and Associations nor the European Commission are liable for any use that may be made of the information contained therein.

Financial & competing interests disclosure

M Morrison is funded by the ESRC project grant ES/P002943/1 “*Biomodifying technologies and experimental space.*” For C George and S Harmon, the research leading to these results has received funding from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement No 115582-EBiSC (European Bank for induced pluripotent Stem Cells); resources of which are composed of financial contribution from the EU’s Seventh Framework Program (FP7/2007–2013) and EFPIA companies in kind contribution. J Kaye is funded by Wellcome Trust award 096599/2/11/Z which includes provision for paying open access fees. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Executive summary

- This paper reviews the impact of changing data protection law for induced pluripotent cells (iPSC) researchers and iPSC-based biobanks and offers recommendations on how to operate in ways that ensure compliance with the requirements of the forthcoming European General Data Protection Regulation (GDPR).

The General Data Protection Regulation

- The European GDPR comes into force in May 2018 and will apply in all EU Member States, in the UK and to any transfer of data derived from EU citizens to territories outside the EU.
- Its purpose is to harmonize the law regarding the protection of personal data and clarify the obligation of any persons working with personal data.
- Data related to health, biometric data and genetic data can all be used to characterize human-derived iPSC and all come under the scope of the Regulation if they are not fully anonymized.

Consent

- Consent forms should collect explicit consent for: donation of biological sample for reprogramming; collection of personal data; collection of genetic and biometric data; and sharing of sample and associated data.
- Academic studies that generate iPSC for the study of patients in a particular diagnostic category should also consider the consent requirements to allow these celllines to be deposited in a biobank for subsequent distribution.

Record keeping & accountability

- Data controllers and data processors must keep adequate and up to date records of what data have been collected, stored and accessed.
- Institutions processing sensitive and personal data on a large-scale to appoint a Data Protection Officer. This will include large biobanks, universities and companies with more than 250 employees.
- It is also critically important to maintain robust institutional connections so that pseudonymized samples and data can be connected back to individual donors.

Data minimization

- The GDPR advocates minimizing the amount of data collected and implementing privacy by design approaches to achieve this.
- Both these objectives are challenging to achieve for biobanks dedicated to making samples and data maximally available to the global research community.

Transfer of samples/data outside the EU

- The GDPR prohibits the transfer of personal data outside the EU unless the foreign jurisdiction is adjudged to have adequate data protection measures in place.
- In the absence of a complete list of approved jurisdictions, the responsibility currently falls on the institution controlling the data to ensure that appropriate safeguards are in place.
- Material Transfer Agreements, Codes of Conduct and other agreements offer some measure of protection.

Discussion

- A table of key data protection points for scientists and biobanks working with human derived iPSC to take in to account is provided.

References

Papers of special note have been highlighted as: • of interest

- 1 Scudeleri M. News feature: a decade of iPSC cells. *Nature* 534, 301–312 (2016).

- 2 Inoue H, Nagata N, Kurokawa H, Yamanaka S. iPS cells: a game changer for future medicine. *EMBO J.* 33(5), 409–417 (2014).
- 3 Kimbrel EA, Lanza R. Pluripotent stem cells: the last 10 years. *Regen. Med.* 11(8), 831–847 (2016).
- 4 Avior Y, Sagi I, Benvenitsky N. Pluripotent stem cells and disease modelling in drug discovery. *Nat. Rev. Mol. Cell Biol.* 17(3), 170–182 (2016).
- 5 Suter-Dick L, Alves PM, Blaauboer BJ *et al.* Stem cell-derived systems in toxicology assessment. *Stem Cells Dev.* 24(11), 1284–1296 (2015).
- 6 Fatehullah A, Tan SH, Barker N. Organoids as an *in vitro* model of human development and disease. *Nat. Cell Biol.* 18(3), 246–254 (2016).
- 7 McKernan R, Watt FM. What is the point of large-scale collections of human induced pluripotent stem cells? *Nat. Biotechnol.* 31(10), 875–877 (2013).
- **Reviews the recent growth in large-scale human-induced pluripotent stem cell banking projects and their aims.**
- 8 De Sousa PA, Steeg R, Wachter E *et al.* Rapid establishment of the European bank for induced pluripotent stem cells (EBiSC) – the hot start experience. *Stem Cell Res.* 20, 105–114 (2017).
- 9 Streeter I, Harrison PW, Faulconbridge A *et al.* The human-induced pluripotent stem cell initiative – data resources for cellular genomics. *Nucleic Acids Res.* 45(D1), D691–D697 (2017).
- 10 Morrison M. Infrastructural expectations: exploring the promise of international large-scale induced pluripotent stem cell banks. *New Genet. Soc.* 36(1), 66–83 (2017).
- 11 Isasi R, Andrews PW, Baltz JM *et al.* Identifiability and privacy in pluripotent stem cell research. *Cell Stem Cell* 14(3), 427–430 (2014).
- **Sets out general data protection responsibilities for human-induced pluripotent stem cell researchers and biobanks.**
- 12 Council Directive (EC) 95/46 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 L281/31 .
- 13 Council Regulation (EC) 2016/679 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 L119/1 .
- 14 Morrison M, Moraia LB, Steele JC. Traceability in stem cell research: from participant sample to induced pluripotent stem cell and back. *Regen. Med.* 11(1), 73–79 (2016).
- 15 Knoppers BM, Isasi R, Benvenisty N *et al.* Publishing SNP genotypes of human embryonic stem cell lines (hESC): a policy statement. *Stem Cell Rev.* 7(3), 482–484 (2011).
- 16 Reichel J. Oversight of EU medical data transfers – an administrative law perspective on cross-border biomedical research administration. *Health Technol.* 27(3), 505 (2017).
- 17 Pouillet Y. EU data protection policy. The Directive 95/46/EC: ten years after. *Comput. Law Security Rev.* 22(3), 206–217 (2006).
- 18 Information Commissioner's Office. What derogations does the GDPR permit? <https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/national-derogations/>
- 19 Denham E. How the ICO will be supporting the implementation of the GDPR. The Information Commissioner's Office blog (2016). <https://iconewsblog.wordpress.com/2016/10/31/how-the-ico-will-be-supporting-the-implementation-of-the-gdpr>
- 20 Harbo T-I. The function of the proportionality principle in EU law. *Eur. L. J.* 16(2), 158–185 (2010).
- 21 Council Directive (EC) 2004/23 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 L102/48.
- 22 Council Directive (EC) 2006/86 24 October 2006 implementing Directive 2004/23/EC of the European Parliament and of the Council as regards traceability requirements, notification of serious adverse reactions and events and certain technical requirements for the coding, processing, preservation, storage and distribution of human tissues and cells, OJ L294/32.
- 23 WHO. WHO guiding principles on human cell, tissue and organ transplantation. *Transplantation* 90(3), 229–233 (2010).
- 24 Lomax GP, Chandros HS, Rosario I. The DISCUSS project: revised points to consider for the derivation of induced pluripotent stem cell lines from previously collected research specimens. *Stem Cells Transl. Med.* 4(2), 123–129 (2015).
- 25 Ogbogu U, Burningham S, Ollenberger A *et al.* Policy recommendations for addressing privacy challenges associated with cell-based research and interventions. *BMC Med. Ethics* 15, 7 (2014).
- **Reports on a multidisciplinary workshop on how to best manage privacy issues associated with cell-based research.**
- 26 *Article 29 Working Party, Guidelines for Identifying a Controller or Processor's Lead Supervisory Authority*, 16/EN WP 243 rev.01.
- 27 Human Pluripotent Stem Cell Registry (hPSCReg). www.hpscereg.eu/
- 28 The Stem cell Database (StemDB). www.stemdb.org
- 29 *BBMRI-ERIC Answers to Frequently Asked Questions Updated Version 2.0 Prepared by the BBMRI-ERIC Common Service ELS* (2017). www.bbMRI-eric.eu/wp-content/uploads/2017/03/BBMRI-ERIC_FAQs_on_the_GDPR_V2.0.pdf
- **Covers commonly asked questions about the forthcoming European General Data Protection Regulation and how it relates to cell and tissue biobanks.**

- 30 Gürses S, Troncoso C, Diaz C. Engineering privacy by design. Paper presented at the 4th Conference on Computers, Privacy and Data Protection. Brussels, Belgium, 25–27 January 2011 (2011).
- 31 Kaye J, Hawkins N. Data sharing policy design for consortia: challenges for sustainability. *Genome Med.* 6(1), 4 (2014).
- 32 European Commission. Model Contracts for the transfer of personal data to third countries. http://ec.europa.eu/justice/data-protection/international-transfers/transfer/index_en.htm
- 33 The Information Commissioner's Office. Model Contract Clauses: International transfers of personal data. https://ico.org.uk/media/1571/model_contract_clauses_international_transfers_of_personal_data.pdf
- 34 Isasi R, Knoppers BM, Lomax G. Sustained interaction: the new normal for stem cell repositories? *Regen. Med.* 6(6), 783–792 (2011).
- 35 Kaye J, Whitley EA, Lund D, Morrison M, Teare H, Melham K. Dynamic consent: a patient interface for twenty-first century research networks. *Eur. J. Hum. Genet.* 23, 141–146 (2014).
- 36 Teare HJA, Morrison M, Whitley EA, Kaye J. Towards 'Engagement 2.0': Insights from a study of dynamic consent with biobank participants. *Digital Health* 1, 1–13 (2015).
- 37 Stephens N, Atkinson P, Glasner P. The UK Stem Cell Bank: securing the past, validating the present, protecting the future. *Sci. Cult.* 17(1), 43–56 (2008).
- 38 Zawati MN, Borry P, Howard C. Closure of population biobanks and direct to consumer genetic testing companies. *Hum. Genet.* 130(3), 425–432 (2011).
- 39 Chalmers D, Nicol D, Kaye J *et al.* Has the biobank bubble burst? Withstanding the challenges for sustainable biobanking in the digital era. *BMC Med. Ethics* 17(1), 39 (2016).
- **Describes four 'waves' of biobank design and governance, discusses what has been learned and sets out future prospects.**
- 40 Kaye J. From single biobanks to international networks: developing e-governance. *Hum. Genet.* 130(3), 377–382 (2011).
- 41 Zika E, Paci D, Braun A *et al.* A European survey on biobanks: trends and issues. *Public Health Genomics* 14(2), 96–103 (2011).