phg foundation
making science work for health

Genomic medicine and research: how does the GDPR apply?

Authors

Colin Mitchell, Johan Ordish and Alison Hall

Acknowledgements

This work was funded by the Information Commissioner's Office Grants Programme

URLs in this report were correct as of July 2020

This report can be downloaded from:
www.phgfoundation.org

Published by PHG Foundation
2 Worts Causeway
Cambridge
CB1 8RN
UK

+44 (0)1223 761900

January 2020

Correspondence to:
intelligence@phgfoundation.org

How to reference this report:
Genomic medicine and research: how does the GDPR apply? (2020)

Summary

This discussion paper addresses how the GDPR applies to personal data in genomic medicine and research. It outlines some challenges in meeting the requirements of the GDPR including: establishing a legal basis for processing, fulfilling conditions for processing of genetic or health data, complying with rights and obligations while data are being processed, and meeting requirements for international data transfers outside the EU/EEA.

Key points

- Processing genetic or health data is not automatically high risk but it requires a high level of protection and safeguards
- Uncertainty and differences in opinion about the requirements of the GDPR could impact data sharing for genomic medicine and research
- Reaching consensus about how the GDPR applies in specific situations will not be easy but a sector-specific code of conduct could help streamline compliance and improve confidence.

Introduction

As explored in our previous discussion paper *Genomic medicine and research: when does the GDPR apply?* the General Data Protection Regulation (GDPR) has updated the legal framework for the processing of personal data across the EU. The Regulation is designed to strengthen data subjects' rights and bring data protection law up to date with the technological and societal changes that have taken place in the two decades since the previous Directive was created. Unlike the Directive, the GDPR is based on the 'fundamental right to the protection of personal data' (Art 16 TFEU) and it provides only limited scope for tailoring of rules by Member States (MSs). There are new rights, such as the right to erasure, new obligations for data controllers and processors, new governance requirements (e.g. for a data protection officer reporting directly to the highest management level) and greatly enhanced potential fines of up to 4% global annual turnover.

The GDPR and the Data Protection Act 2018:

Unlike the previous directive-based regime for data protection, the GDPR applies directly in all Member States and is the direct source of most data protection principles rights and obligations in the UK.

The UK Data Protection Act (DPA) 2018 applies, supplements or varies aspects of this law that Member States (MS) have been expressly allowed to tailor.

In this discussion paper we focus on how the GDPR impacts the use of genetic or genomic data (and associated health data) in healthcare and research. These impacts include challenges in transparently establishing a legal basis for processing, processing 'special category' data (including genetic data), meeting obligations, enabling data subjects' rights and demonstrating a lawful basis for international data transfers.

Principles for processing

Under the GDPR it is important to note that there are seven overarching principles for data processing (Art 5) that must be met when processing personal data in genomic healthcare or research. A failure to comply with them may result in significant sanctions.[1] The principles remain broadly the same in the GDPR as the Data Protection Directive with a few changes and additions.

They are that:

- Personal data shall be processed lawfully, fairly, and transparently
- Processed for specific purposes and not further processed in a manner that is incompatible with those purposes (purpose limitation principle)
- Limited to what is necessary in relation to those purposes (data minimisation)
- Data shall be accurate and kept up to date (accuracy)

- Kept in an identifiable form for no longer than is necessary (storage limitation)
- Processed in a way that protects them from unauthorised processing, loss, destruction or damage (integrity and confidentiality)
- Data controllers are responsible for demonstrating compliance with these principles (accountability)

In the genomic context, there are some particular challenges, for example, ensuring sufficient transparency in complex data initiatives, or, determining how accuracy applies to genetic results. There is some flexibility to purpose limitation for genomic research, which is deemed not to be incompatible with the initial purposes of processing so long as it is in accordance with Art 89 of the GDPR (discussed further below). This means that the secondary use of genetic data for research purposes will not necessarily require a new legal basis, but the principles of fairness and transparency will still be important, and new information may need to be given to data subjects (see the right to information below).[2] The principles also provide an important aid in the interpretation of more specific obligations within the GDPR, including the requirement for a specific legal basis for processing.

Establishing a legal basis for processing

Although the GDPR carries over many of the requirements of the previous Data Protection Directive, it still gives rise to some challenges when establishing a legal basis for processing of personal data, as required by Art 6. In particular, the GDPR has introduced new, more stringent standards for consent leading some data controllers to consider alternative legal bases for processing of genetic or health data. The choice of legal basis is also important because different rights, or exceptions to them, flow from each choice of legal basis. Under the GDPR, personal data shall only be processed if one of six legal bases can be satisfied (Art 6 (1)). The requirements for consent are discussed below, followed by brief consideration of the alternative legal bases.

The challenge of consent

Article 4 (11) defines the 'consent' of the data subject as:

> Any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action signifies agreement to the processing of personal data relating to him or her

Art 7 sets out further conditions, including that the data subject must be informed that they have the right to withdraw their consent at any time (Art 7(3)), and the background text (recitals) provide further guidance. Recital 42 explains that at a minimum a data subject should be informed of the identity of the controller and the purposes of the processing, and that there must be a genuine free choice to refuse or withdraw consent without detriment. Recital 32 explains that consent should be given by a clear, unambiguous affirmative act (excluding silent or opt-out consent) and that consent should cover all processing activities carried out for the same purposes.

One of the challenges under the GDPR is that consent is not valid where there is a 'clear imbalance' between the data subject and the controller, in particular where the controller is a public authority (recital 43), which is frequently the case in healthcare and research. The European Data Protection Board (EDPB) – the independent body which ensures consistency of data protection rules across the EU – has cautioned that where a participant is not in good health, or when they belong to economically or socially disadvantaged groups, there is likely to be an imbalance of power.[3]

In addition, a particular challenge for genomic research is the ambiguity around how specific consent must be. In general, it is clear that consent is lawful only if given for one or more 'specific purposes' (Art. 6 (1)(a)) and it should allow for separate choices to be given to different 'personal data processing operations' (recital 43). This is explained by the Article 29 Working Party (the predecessor to the EDPB) as requiring 'granularity', so where appropriate, separate consent should be obtained for different purposes.

An important question for genomic research, particularly long-term big data research, is how specifically defined the research purposes must be at the outset. The Article 29 Working Party previously explained that enough information should be provided about specific purposes for the data subject to understand the implications of their choice,[4] and to assess whether the law and safeguards have been complied with.[5] They concluded that a purpose which is 'vague or general, such as … "future research" … will – without more detail – usually not meet the criteria of being "specific".'[6]

However, in recital 33 of the GDPR it is acknowledged that:

> It is often not possible to fully identify the purposes of personal data processing for scientific research purposes at the time of collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research.

This suggests that a broader consent may be possible for processing of genomic data in 'areas' of scientific research purposes, subject to ethical oversight. Such flexibility would align with other aspects of the GPDR, such as the explicit exemption from the purpose limitation principle for scientific research. However, the Article 29 Working Party also cautions that the requirement for specific consent will apply unless it is not possible to sufficiently specify the purposes for data processing at the outset, providing an 'exception that the purpose may be described at a more general level.'[7]

The Article 29 Working Party also state that the 'flexible approach' of recital 33 will be subject to a higher degree of scrutiny when special categories of data, such as genetic and health data (see below), are processed on the basis of explicit consent.[8] Furthermore, a controller 'must seek other ways to ensure the essence of the consent requirements are served best', such as seeking consent to each defined stage of research as it progresses, or providing regular updates on the development of research purposes.[9] This could be read as an encouragement for dynamic consent if purposes cannot be sufficiently specified at the outset.

The potential power imbalance between researchers and participants, and the challenge of ensuring sufficiently specific consent, have contributed to recommendations against consent as a legal basis for scientific research processing. The NHS Health Research Authority have stated that 'the legal basis for processing data for health and social care research should not be consent',[10] for both public authorities and commercial companies or charitable research organisations. As we discuss further below, if processing is based on consent, a number of rights and obligations follow, including the right to withdraw consent at any time, which can be challenging for some genomic research projects.

Even if consent is not chosen as a legal basis under the GDPR for processing personal data, consent may still be required in order to comply with other legal and ethical standards, including the common law of confidentiality. This can make explaining to individuals how their data will be used and managing their expectations difficult, especially that withdrawal of consent may not prevent further processing for research purposes (see below). This is a challenge for those who are seeking to be clear and transparent about processing.

Alternative legal bases

There are five other legal bases available but distinctions between public authorities and private organisations influence which are most appropriate. For public authorities, the most appropriate legal basis is often likely to be 'performance of a task carried out in the public interest' (Art 6(1)(e)). It will apply if the processing is necessary to perform a task, or in the exercise of official authority, function or power, which has a clear basis in law.[11] This basis can only be relied upon by public authorities in the UK subject to freedom of information legislation or otherwise specified in regulations.[12] It is likely to be the most useful legal basis for public sector healthcare or research organisations processing genetic data and it will clearly apply to provision of healthcare by the NHS.[13] It should also be available to a university undertaking processing of personal data for medical research purposes.[14] However, this basis is not available to private or charitable organisations. For them, the legal basis of 'legitimate interests' (Art 6 (1) (d)) is likely to be the best option in many cases.

Legitimate interests are the most flexible basis for processing personal data (although they are no longer available as a legal basis for processing by public authorities performing their tasks or exercising their official authority). The concept of a legitimate interest is very broad and under the GDPR has been extended to include the legitimate interests of any third party, not only recipients of data. The Information Commissioner's Office (ICO) guidance is that it may extend to even trivial or controversial interests but that vague, unethical and unlawful interests would not count as legitimate.[15] However, data processing must be necessary for the purposes of those legitimate interests, which means it must be a reasonable and proportionate way to achieve those ends, and may be outweighed by the impact of processing on the interests or fundamental rights and freedoms of the data subject.

The ICO recommend that controllers perform a 'legitimate interests assessment' (LIA) to assess the balance of the controller's interest and the impact on the data subject, and to record the justification for processing. The reasonable expectations of the data subject are an important factor in this assessment. In the genomics context, evidence of patient or participant expectations could provide support for processing. The sensitivity of genetic or health data and the need for heightened protection of such data (especially of children's data) are significant but not insurmountable barriers to processing on the basis of legitimate interests.

Other legal bases may apply in specific contexts. For example, processing is lawful if it is necessary for the 'performance of a contract' (Art 6 (1)(b)). This could be the case in private healthcare or direct to consumer genetic services. However, processing must be necessary for performance of (or entry into) the contract, so if the collection of valuable personal data is just a part of the business model another legal basis will be required.[16]

Some forms of processing of genetic/health data may be necessary for 'compliance with a legal obligation' (Art6 (1)(c)). This only applies if processing of personal data is a reasonable and proportionate way of complying with a clear statutory or common law obligation (Art 6(1)(c)). For example, sharing healthcare data via a statutory gateway with bodies such as NHS Digital. It could perhaps provide a basis for disclosure of genetic information to relatives if a clear legal duty to disclose genetic information to relatives of the data subject is established in UK law.[17]

As a matter of last resort, processing will be lawful if it is necessary to protect the vital interests (generally interpreted as matters of life and death[18]) of the data subject or another person (Art 6(1)(d)). This is unlikely to be met for most genetic tests unless the threat of death is significant and proximate, for instance, where using whole genome sequencing to inform the diagnosis of rare genetic diseases in severely ill babies and young children. It will not apply if processing can be justified by an alternative basis (recital 46).

Summary

Overall, the challenges with consent mean that for many uses of health and genomic data other legal bases will often be more appropriate, particularly for research purposes. There are obvious alternatives, public task for public authorities such as university medical researchers, and legitimate interests for private or charitable organisations. However, in the context of genomic data the legitimate interests of the processor need to be balanced sensitively with the rights and interests of the data subject and even if consent is not the legal basis for processing it will often be required for other legal and ethical reasons. Communicating these distinctions to data subjects may be challenging.

Conditions for processing genetic and health data

Even if a legal basis has been established for processing personal data, the processing of genetic data, data concerning health and biometric data (amongst other 'special categories') is actually prohibited by the GDPR unless one of the ten conditions in Art 9 (2) apply. We discuss some of these special categories in our previous discussion paper, and it should be noted that they are potentially very broad: 'data concerning health' includes all data which reveal information about the health status of an identifiable or identified individual, 'genetic data' extends to data which give unique information about the physiology of an individual, and 'biometric data' includes facial images and other images which allow the unique identification of an individual. This means that most of the individual-level data that result from genetic analysis and frequently accompany genetic results as phenotypic data will fall within these special categories of data, provided that they are reasonably likely to identify an individual.

Which conditions?

In the genetic or genomic context some of the Art 9 (2) conditions are particularly relevant. These include that processing is necessary for medical purposes, public health purposes or for scientific research. However, each condition has its own requirements and consequences for further obligations and data subjects' rights. We briefly outline these requirements for the most relevant conditions and highlight the potential challenge for cross-border genetic or genomic initiatives complying with Art 9.

Medical or public health purposes

Genetic and health data may be processed for medical or public health purposes: Art 9 (2)(h) allows processing for 'the purposes of preventive or occupational medicine … medical diagnosis, the provision of health or social care or treatment or the management of health or social care services on the basis of Union or Member State law or pursuant to contract with a health professional'. The DPA 2018 provides a basis in UK law[19] and implements the 'secrecy' safeguards required by the GDPR (Art 9 (3)) so that processing must be under the responsibility of a health or social work professional or another person who in the circumstances owes a duty of confidentiality.[20] If there is a public health objective to processing data – such as separating human DNA from virus DNA as part of pathogen sequencing during an outbreak[21] – the Art 9(2)(i) condition will be more suitable, which applies if 'processing is necessary for reasons of public interest in the area of public health'. This includes ensuring high standards of quality and safety of health care, medicines or medical devices, and must be supervised by a professional under a duty of confidentiality.[22]

Scientific research purposes

An important addition to the GDPR is that special category data may be processed if necessary for scientific research purposes (Art 9 (2)(j)). This requires implementation in national law, which the UK has done with some further conditions: Processing must be in the public interest,[23] and not 'likely to cause substantial damage or substantial distress to a data subject'.[24] If processing is carried out for the purposes of 'measures or decisions with respect to a particular data subject' it must have been approved by a research ethics committee.[25] Researchers must also put in place 'appropriate safeguards' for the rights and freedoms of the data subject (Art 89 (1)). In line with the broader principle of data minimisation, if it is possible to achieve the research purposes by further processing which 'does not permit or no longer permits the identification of data subjects', Art 89 requires that this should be done. Overall, to meet this condition genetic or genomic data research must be as limited as possible and include technical and organisational safeguards such as pseudonymisation. If individual measures, for example recontact with results, are anticipated, research must have REC approval. It may be challenging to rely on this basis for cross-border research if Member States implement different requirements (or even if some do not enact this derogation at all). However, research processing has the benefit of a range of exemptions from some of the rights and obligations in the GDPR (see below), which could considerably reduce the burden on researchers.

It is also possible to process genetic, health and other special category data on the basis of consent but this must meet all the standards discussed above and must also be 'explicit' (Art 9 (2)(a)). This requires an express statement of consent, for example in a written statement signed by the data subject or a similar positive action using an electronic form.[26] Although the challenges discussed above mean that alternative derogations are recommended as preferable to consent for purposes like scientific research, consent may still be the best available option if organisations are subject to multiple overlapping Member State laws, and research provisions have not been implemented in all of them.

Other conditions

Some other Art 9 (2) conditions may apply to particular uses of genomic data. For example, special category data may be processed if the processing relates to personal data which are manifestly made public by the data subject (Art 9(2)(e)). It is possible that this could apply if individuals have consciously published their genetic or health information on an open website, for example on an ancestry database. However, this derogation should be interpreted in line with the higher protection required for special category data and the data subject herself must have positively and knowingly made data public, as opposed to accessible within a limited group of people, even if the data have become publicly accessible thereafter.[27]

Special category data may also be processed if it is necessary to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent (Art 9 (2)(c)). As previously discussed in relation to legal bases, this is a matter of last resort where there is a risk of life and death, and where alternatives are unavailable, so it will not apply in many circumstances. A further option that could become relevant to the genomic context, is if processing is necessary for reasons of substantial public interest based in EU or UK law (Art 9 (2) (g)). The DPA 2018 sets out specific circumstances where this condition may be met in the UK law.[28] None are currently directly relevant to healthcare, public health or research purposes but there is scope for this list to be amended over time.

National variations

One challenge for cross-border genomic medicine and research initiatives is that though the GDPR harmonises many aspects of data protection law, it also provides considerable scope for national variations for processing special category data. Article 9 (4) explicitly allows Member States to introduce (or maintain) further restrictions, conditions or limitations to the processing of genetic, health and biometric data. The UK has yet to implement specific conditions but other countries such as France have.[29] Some nations have gone further including Finland and Italy, by introducing specific requirements for the processing of genetic data.[30] Moreover, because most of the Art 9 conditions also require implementation or authorisation in EU or Member State law, some, e.g. for scientific research, may not be available in all jurisdictions. This could mean explicit consent becomes the only realistic cross-border option.[31] Finally, when Member States provide a legal basis for derogations like scientific research purposes, they must provide 'suitable and specific measures to safeguard the fundamental rights and interests of the data subject' (Art 9(2)(j)). The choice of these measures is left to the Member States so significant differences arise.

For example, Ireland has implemented measures which require explicit consent in health research unless a committee can be satisfied that the public interest in the research 'significantly outweighs' the public interest in requiring the explicit consent of the data subject.[32] This means that cross-border genetic or genomics projects are likely to face a complicated regulatory environment and will have to ensure that their approvals, policies, processes and patient/participant information meet the requirements of each relevant jurisdiction. This could mean that despite the challenges already discussed (and noting in particular different national conditions and safeguards relating to genetic data), explicit consent for the processing of special category data may be the preferred option for those seeking a single streamlined cross-border approach.

Fulfilling data subject rights

Another area of complexity and potential challenge in the context of genomic data relates to the range of data subject rights that may apply, and the systems and processes that will be required to support the exercise of them in healthcare or research contexts. Some are new rights, such as the right to be forgotten (Art 17) whereas others carry over from the Directive but may be enhanced, such as the right to information which is bolstered to support transparency. Each right raises issues for this context so we will discuss them briefly with a focus on the aspects that are most relevant to genomic data processing.

Determining which rights apply and when?

There are an array of exceptions to many data subject rights depending on the nature of the processing and the choice of legal basis. A potentially important restriction for genomic contexts is found in Article 11 which restricts the rights under Arts 15-20, (see below) if the data controller can show they are 'not in a position to identify the data subject' (unless the data subject provides further information to facilitate identification). For example, if the data have been pseudonymised and the controller no longer has access to the 'key' then this restriction could apply. As such, Art 11 could provide an important reduction in the burden on data controllers who have taken considerable effort to minimise and de-identify data. It doesn't require the same degree of de-identification as anonymisation because it is only focused on the ability of the controller (not third parties as well) to identify an individual. However, if it is possible – even by significant effort – to identify an individual, for example by obtaining a 'key' from a collaborator, then a controller will not be able to rely on Art 11.

Secondly, some rights are inherently limited according to the chosen legal basis Thirdly, and specifically for scientific research, there are a range of limitations that apply where data are processed in accordance with Art 89 (1), both in the text of the rights themselves and as limited by Member State law (e.g. the UK Data Protection Act 2018) in accordance with Article 89(2). For instance, in the UK the operation of Article 15(1) to (3), (confirmation of processing, access to data and safeguards for third country transfers) Article 16 (right to rectification), Article 18(1) (restriction of processing) and Article 21(1) (objections to processing) are restricted under certain circumstances.[33] This is only to the extent that the application of these provisions would 'prevent or seriously impair the achievement of the [research or statistical] purposes in question'[34] so the onus is on researchers to explain why these rights do not apply.

Finally, Member States are allowed to introduce further restrictions to the operation of data subjects' rights under Article 23 for specific purposes, for example to safeguard the rights and freedoms of others. The rights and the relevant restrictions likely to apply when processing data for healthcare or research purposes are discussed below.

Right to information (Arts 13 & 14)

The rights to information in articles 13 and 14 have the fewest exceptions. A data controller is obliged to provide data subjects with a range of information about how and why data are being processed. This is to support transparency and fairness and so that data subjects can scrutinise and challenge the use of their data, as well as enabling them to secure their rights. It applies both when data have been collected from the data subject (Art 13) and when personal data have been obtained from another source, even publicly accessible sources (Art 14). These rights apply even where other obligations or principles have been limited. For example, although scientific research purposes are deemed not to be an incompatible purpose requiring a new legal basis (see above) the data controller is still obliged to provide information about that new purpose prior to processing (Art 13(3) & 14(4)).

When is information required?

Information is required either at the time of data collection of data from the data subject (Art 13), or if the data was obtained from another (even public) source (Art 14) then:

- Within 'a reasonable period' after obtaining the personal data (not exceeding one month)
- When the personal data is used to communicate with the data subject (e.g. on recontact with new results)
- Where disclosure to another recipient is envisaged
- Where the controller intends to further process the data for another purpose e.g. for scientific research purposes (Arts 13(3) & 14(4))

Controllers must provide details including the key information in the box below. There are some potential challenges in the context of genomic medicine and research. One is that it is not clear how specifically the 'categories' of data recipients must be defined. This could be important for information about research projects. For example, when drawing on the principles of transparency, fairness and purpose limitation, it seems likely that simply describing 'researchers' as a category would be insufficient. At the very least, it seems likely that the type of recipients will be required, for example whether they are public sector or commercial organisations.

However, for scientific research there is a significant exception to the right to information if the data was not obtained from the data subject, and if 'the provision of such information proves impossible or would involve disproportionate effort… or in so far as [this would] render impossible or seriously impair the achievement of the objectives of that processing' (Art 14 (5) (b)). This requires researchers to show that providing information would involve a level of effort and resources that would detract from the research objectives. If so, they should try to use alternative measures, such as making the information publicly available (Art 14 (5) (b)).

What information is required?

- The purposes of the processing and the legal basis for each set of purposes
- The categories of personal data involved
- Any (categories of) recipients of data If applicable the basis on which data are being transferred outside the EU and safeguards involved
- The legitimate interests being pursued if that is the legal basis for processing
- The existence of the other data subject rights
- The right to complain to a supervisory authority
- The existence of automated decision making as referred to in Art 22 (1) & (4)
- The right to withdraw consent if consent is the legal basis for processing

Right of Access (Art 15)

The right of access is a right to obtain a copy of the personal data undergoing processing from the data controller (Art 15 (3)). Despite reports of one data protection authority interpreting this as more of a 'summary' than a complete copy,[35] 'copy' is more generally understood to mean all the data that relate to a specific individual.

This is a major challenge in the genomic context, because a copy of personal data could extend to the full sequence of data resulting from whole-exome or whole-genome sequencing, in addition to the associated clinical or phenotypic information. Such volume of data means that facilitating access is a significant challenge, and it may come as a surprise to patients/participants to discover the volume of data involved. The GDPR also stipulates that if the request is made electronically, a copy should be provided in a 'commonly used electronic form'. What this entails for genetic data is not precisely clear but at least a physical printout of data is not necessary following such a request. However, this is not the same as the requirements of the right to data portability which mandates structured machine-readable formats (see below).

One limitation to the right of access particularly relevant to genetic information is that it should not adversely affect the rights and freedoms of others (Art 15 (4)) and Member States are entitled to restrict the right of access for this purpose (Art 23 (1)(i)).

To this end, the UK has implemented a clarification that data controllers are not obliged to disclose personal data under Art 15, 'to the extent that doing so would involve disclosing information relating to another individual who can be identified from the information',[36] unless the other individual has consented[37] or 'it is reasonable to disclose the information to the data subject without the consent of the individual'.[38] A range of considerations are mentioned as relevant to whether the disclosure is reasonable (e.g. any express refusal of consent by the other individual) so this decision is not entirely at the discretion of the controller.

Interestingly, the data controller is not asked to determine whether the information is 'personal data' relating to the other individual using the standard GDPR test but instead must conduct a slightly different analysis to determine if it is 'information relating to another individual':[39] data which are either directly identifying or which could indirectly identify an individual from 'information that the controller reasonably believes the data subject is likely to possess or obtain.'[40] This could apply to genetic information of familial relevance so a controller faced with an access request for genetic or clinical data is required to focus on what other information the data subject is likely to possess or obtain which could identify a family member.

For scientific research conducted in accordance with Article 89(1) and s19 of the DPA, the right of access will not apply to the extent that it would prevent or seriously impair the achievement of the research purposes, so long as no research results or statistics are published in an identifiable form.[41]

Right to Rectification (Art 16)

Another right that raises some specific questions for genetic or genomic data is the right to rectification of inaccurate data (or, 'taking into account the purposes of processing', the completion of incomplete data). This is an important corollary to the principle of data accuracy (that data should be kept accurate and up to date) and raises the question of when genetic or genomic data can be considered inaccurate. The GDPR doesn't define accurate or inaccurate but the DPA 2018 does describe inaccurate in relation to personal data as meaning 'incorrect or misleading as to any matter of fact'. (DPA 2018, s205 (1))

In the genetic context, data such as variant classifications may become out of date. Does this mean that all records need updating as a matter of course, to correct inaccurate or misleading classifications? Part of the answer is that data are only required to be accurate on their own terms. For example, if they are the results of a test with a significant margin of error, so long as that margin of error is explained, the data will be accurate even if they may be erroneous. If there is no reference to the chance of error then the data may be misleading. Equally, if results are reported as accurate according to the current state of the evidence, this will not be misleading. If a conclusion about a genetic result is a matter of opinion, this should also be explained. Even if data are updated, it could be that earlier 'inaccuracies' should be retained as an accurate record of the analytical or decision-making process.[42]

In the longer term, it is an open question whether the right to rectification could be applied to require the updating of records if their results are clearly no longer accurate. On the one hand it could be argued that the records are an accurate account of the results at the time of testing and analysis. On the other hand, if the results potentially have influence on the data subject's eligibility for screening or have an impact in other ways (perhaps even on insurance options in the future), then arguably, a failure to rectify results could result in records that are factually inaccurate or misleading about the data subject's true health status.

As with other rights, in scientific research the right to rectification is limited in the UK to the extent that it would 'prevent or seriously impair the achievement of the [research or statistical] purposes in question.'[43] This means the issue of data accuracy and rectification is likely to be a greater challenge for processing in clinical activities than scientific research.

Right to erasure or to be forgotten (Art 17)

The right to erasure is a new addition in the GDPR with potentially far reaching impact, particularly if it can be applied to research databases. However, this right is largely restricted to circumstances where data are not processed for health, public health[44] or research purposes (if it would seriously impair the research)[45] and when processing is based on consent. If consent is relied on as a legal basis for processing, or if explicit consent is used to justify the processing of special category data, the right to erasure should be complied with unless there are other legal grounds for processing.[46] If a request for erasure is valid but the data has already been made public, the data controller is obliged to inform other controllers who are processing the data about the request.[47] This is particularly important if the data have been transferred internationally.[48] An ambiguity that could have a significant impact on further genetic research is whether erasure can be validly achieved through anonymisation or if it requires complete deletion of data. Whilst some authorities agree that anonymisation could suffice,[49] others disagree and argue that this disempowers data subjects who may wish to prevent the future use of that data.[50] In the genomics context, this could mean the difference between requiring efforts to remove data from a database such as ClinVar, or, accepting that 'anonymised' data can still be used for research.

Right to data portability (Art 20)

Another new addition in the GDPR is the right to data portability. A more limited form of the right to access personal data, portability is aimed at ensuring data are provided in 'commonly used and machine-readable' formats and that a data controller doesn't hinder the transmission of personal data to another controller. This right is quite limited. It only applies to data that the data subject 'has provided to a controller', not data which have been obtained from a third party or which the data controller has generated themselves. This means that it applies to the 'input' data provided by the data subject and not the results of further analyses. Guidance from the Article 29 Working Party contrasts data provided by the data subject with 'inferred data' and 'derived data' created by the data controller through analysis of data 'provided by the data subject'.[51]

Given that sequencing data and results are derived from an analysis of the material provided by the data subject, genetic data and results are not likely to fall within the right to data portability.[52] Moreover, this right only applies when processing is on the basis of consent or contract, where special category processing is justified via explicit consent, or where processing is 'carried out by automated means' (Art 20 (1)). Even if consent were the basis or justification for processing genetic data, neither the results of sequencing or further analysis need to be provided in machine readable format, because they are not the data which the data subject provided.

Right to object (Art 21)

The right to object may apply to any processing of genetic or health data (including profiling) based on legitimate interests or performance of a public task.[53] If a data subject objects, the controller may only continue processing if they demonstrate compelling legitimate grounds which override the rights and freedoms of the data subject. However, there is a specific exemption for scientific research conducted on the basis of the performance of a task in the public interest (Art 21 (6)), so public research institutions will not have to comply with an objection. This means it is likely to only be private or charitable research bodies who have to demonstrate compelling grounds to continue processing. How 'compelling' is to be assessed is unclear.

Right not to be subject to solely automated processing (Art 22)

A final right that applies only in narrow circumstances is the right 'not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.' This will apply unless the processing is based on explicit consent (or contract, if the data are not special category data), in which case there should at least be the right to obtain human intervention, express a point of view or contest the decision.* At present this is highly unlikely to apply in the health or genomic context, even with advanced AI-based processing, because there will almost certainly be a human health care professional who can intervene in the decision-making process if a decision has a serious and significant impact. However, the principles of fairness, accountability and transparency all suggest that information may need to be provided (depending on the context) about how data are being processed using AI and other automated processing, even if the decision is not solely automated.

* Or if it is in the substantial public interest on the basis of Member State law (Art 22 (4)) and safeguards are in place

Privacy and security of genetic data

Beyond data subject rights, the GDPR also requires data processors and controllers to comply with a range of further obligations, depending on the scope, context and purposes of processing, and the potential for harm. In particular, many obligations require proportionate responses to the sensitivity and risk of processing and these are likely to be elevated in the processing of genetic or genomic data, particularly on a large scale.

Article 24 requires data controllers to implement 'appropriate technical and organisational measures' according to the nature, scope and purposes of processing. Similarly, Article 25 requires 'data protection by design and default' which means a context-sensitive and proportionate implementation of safeguards, such as pseudonymisation, to implement data protection principles and protect the rights of data subjects.

Article 32 adds further detail on the measures that are required to ensure security of processing and, where processing 'is likely to result in a high risk to the rights and freedoms of natural persons', a controller must carry out a data protection impact assessment (DPIA) in accordance with Art 35.

Although genetic data are particularly sensitive, the EDPB have made clear that the processing of genetic information does not automatically count as high risk in itself.[55] The Art 29 Working Party guidance, endorsed by the EDPB, is that the processing of sensitive data is one criterion that could lead to a DPIA being required if another criterion is also met. These include large scale processing of data concerning vulnerable subjects (as in the case of patients or children), or, if the processing involves evaluation and scoring (including profiling and predicting) of data subjects' health.[56] It is likely that the processing of genetic data will require a DPIA in many circumstances including in research projects which process large amounts of personal data, and the implementation of significant technical and organisational safeguards.

Data sharing and international transfers

Some of the major concerns caused by the GDPR relate to its impact on genomic data sharing. In part, these result from specific provisions for transfer of personal data outside the EU/EEA, which are not necessarily working for all international genomic collaborations. Another aspect is the challenge of reaching agreement amongst international partners about compliance, even within the EU. It is possible that a code of conduct could streamline compliance and reach agreement on best practice, and help facilitate the flow of genetic and health data.

International transfers

As well as the obligations mentioned above, including the need for a lawful basis for processing, specific provisions apply to transfer of personal data outside the EU/EEA. provisions apply to transfer of personal data outside the EU/EEA. Under Chapter V of the GDPR this may only take place where the European Commission has determined there is an adequate level of protection (so called 'adequacy decisions', Art 45); where there are appropriate safeguards (Art 46); or as a last resort, where an exception such as explicit consent, applies (Art 49). The number of countries/sectors which have been approved as ensuring an adequate level of protection is small: 'adequacy decisions' have been adopted for 11 countries,[57] the commercial sector in Canada, and companies certified in the USA under Privacy Shield. For data transfers outside these adequacy decisions (for example, international genomics collaborations involving university researchers) another mechanism is required to legitimate a transfer. The most obvious of which are the standard data protection clauses that have been approved by the Commission as safeguards under Art 46. However, there have been reports of challenges with these standard clauses because the clauses on liability, jurisdiction and governing law may conflict with the local laws of third country public institutions. Because all contract terms have to be included to comply with Art 46, these conflicts cannot easily be resolved.

An alternative that has been welcomed by biomedical researchers is the potential for an approved code of conduct to act as a safeguard under Article 46 for international transfer. For example, the Biobanking and BioMolecular resources Research Infrastructure (BBMRI-ERIC) and collaborators are working on a code of conduct for health research[58] which, if approved by the EDPB and Commission, could provide an adequate safeguard for international data transfer. However, no such code has yet to be approved by the Commission and achieving this agreement among partners and with the EDPB may be challenging. Furthermore, an expert but independent body must be accredited by the supervisory authority to monitor compliance by controllers and processors (including those in third countries) who adhere to the code (Art 41). These requirements suggest that developing an approved code of conduct, even if sector-specific, will take considerable time and effort, and require continuing oversight after development.

If no adequacy decision is in place and none of the safeguards are appropriate then an international data transfer may take place subject to the conditions in Art 49 (1). These conditions include: that the data subject has provided specific and explicit consent to the transfer, having been warned of the specific risks of the transfer in the absence of an adequacy decision or safeguards under Art 46 (this is in addition to consent obtained under Arts 6 or 9 if they are the basis for processing personal or special category data); that the transfer is necessary in order to protect the vital interests of the data subject or of other persons, where the data subject is physically or legally incapable of giving consent; or as a very last resort, a limited, not repetitive transfer relating to a small number of data subjects may be based on compelling legitimate interests provided they are not overridden by the interests or rights and freedoms of the data subject.

It is also possible for Member States to introduce provisions for transfer on the basis of important reasons of public interest but all Art 49 conditions will be interpreted restrictively, as exceptions from the rule that personal data should only be transferred to a third country with an adequate level of data protection or appropriate safeguards. They cannot occur regularly within a stable relationship, or when a data importer is granted access to a database on a general basis.[59]

Data sharing within the EU/EEA

It is not only international data sharing that may be challenging under the GDPR. Data sharing within the EU or even within a Member State, requires agreement and clarity between partners to ensure that all the requirements of the GDPR are met in relation to processing. For instance, even if data are considered sufficiently anonymised to fall outside the scope of the GDPR, this must be accepted by all parties and there may be differences of opinion. More generally, there are many aspects of the GDPR which Member States may tailor in their own laws. For example, Irish health research regulations chiefly require explicit consent as a specific measure for health research.[60] Although there is a mechanism to try and ensure consistency among data protection supervisory authorities centred on the EDPB, it is likely that different interpretations will continue to arise. As we discuss in relation to special category data above, this means that cross-border sharing of genetic data is likely to remain challenging.

Conclusions and key messages

Once data have been determined to be 'personal data' there are four broad challenges of how the GDPR impacts genomic medicine and research.

Firstly there is the challenge of choosing an appropriate legal basis, in particular, meeting the standards of consent or transparently explaining an alternative in genomic research.

Secondly the challenge of meeting Art 9 conditions for processing special category data, such as health or genetic data, particularly in cross-border initiatives where the law differs across Member States.

The third challenge is understanding and facilitating data subject rights, which vary depending on the purposes of processing or choice of legal basis. There are some particular questions in the genomic context including how the 'accuracy' of genetic results should be assessed.

Finally the GDPR impacts genomic data sharing, both within the EU and internationally. For some initiatives, identifying an appropriate mechanism to legalise the transfer of genetic data outside the EU is a challenge. More fundamentally, it is difficult to reach agreement among partners about the status of data and how best to comply with the rights and obligations required by the GDPR.

At such an early stage in its application, much of the impact that the GDPR has had on the sector is the result of understandable uncertainty about the correct interpretation of its requirements. However, there is a risk that this uncertainty could limit data sharing, in particular cross-border or international genomic data sharing, without proper consideration of how the GDPR applies in the genomics context. This could have a real impact on research and healthcare in the UK and Europe.

Our project, Data protection and genomic data aims to identify priorities for clarification and mechanisms, such as codes of conduct or technical and legal measures, that may help streamline, and improve confidence about, compliance with the GDPR.

Key messages

- As Member States are allowed to tailor the rules governing genetic and health data there is likely to be significant regulatory divergence across Europe for processing these data
- The burden of the GDPR is reduced for scientific research under UK law
- Processing genetic or health data is not automatically high risk but it requires a high level of protection and safeguards
- Consent is challenging even if it is not chosen as a legal basis for processing, as it may be difficult to communicate this to data subjects in ways that are both transparent and accessible
- Determining which rights apply and how in specific contexts is quite complex, and can require considerable resources to put systems and processes in place to facilitate data subject rights.

- Art 11 may reduce the burden of facilitating some data subject rights where data have been so well de-identified that the data controller is no longer in a position to re-identify an individual
- Uncertainty about how to comply with the GDPR gives rise to the danger of a chilling effect on cross border and international data sharing
- A code of conduct under Art 40 of the GDPR could provide a mechanism for streamlining compliance and developing confidence across a sector but this will require considerable effort and agreement to develop

References

1. DLA Piper. Germany: Berlin Data Protection Authority imposes eur 14.5 million fine for 'data cemetary'. Privacy Matters; 6th Nov 2019.
2. The Information Commissioner's Office. Guide to the General Data Protection Regulation (GDPR). ICO website.
3. European Data Protection Board. Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR) (art. 70.1.b)). Adopted on 23 January 2019. Para 20.
4. Article 29 Data Protection Working Party. Guidelines on consent under Regulation 2016/679 as last Revised and Adopted 10 April 2018. p12.
5. Article 29 Data Protection Working Party. Opinion 03/2013 on purpose limitation. Adopted 2 April 2013. pp15-16.
6. Ibid.
7. Article 29 Data Protection Working Party. Guidelines on consent under Regulation 2016/679 as last Revised and Adopted 10 April 2018. p28.
8. Ibid.
9. Ibid: p29.
10. NHS Health Research Authority. Consent in research.
11. The Information Commissioner's Office. Guide to the General Data Protection Regulation (GDPR). ICO.
12. Data Protection Act (DPA) 2018. Section 7.
13. Information Governance Alliance. The General Data Protection Regulation: Guidance on Lawful Processing. 2018.
14. Department for Digital, Culture, Media and Sport and the Home Office. Explanatory Notes to the Data Protection Act 2018. para 85.
15. The Information Commissioner's Office. Guide to the General Data Protection Regulation (GDPR). ICO.
16. European Data Protection Board. Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects. Version 2.0; 8 October 2019. para 37.
17. *ABC v. St George's Healthcare NHS Trust & Ors* [2017] EWCA Civ 336.
18. The Information Commissioner's Office. Guide to the General Data Protection Regulation (GDPR). ICO.
19. DPA 2018. Section 10 (2) and Part 1 of Sch 1.
20. DPA 2018. Section 11 (1).
21. Luheshi L, Raza S, Moorthie S, *et al*. Pathogen Genomics Into Practice. PHG Foundation; 2015.
22. DPA 2018. Section 10 (2) and Sch 1, para 3.
23. DPA 2018. Section 10 (2) and Sch 1, para 4.
24. DPA 2018. Section 19 (2).
25. DPA 2018. Section 19 (3) & (4).

26. Article 29 Data Protection Working Party. Guidelines on consent under Regulation 2016/679 as last Revised and Adopted 10 April 2018. p 19
27. Article 29 Data Protection Working Party. Opinion on some key issues of the Law Enforcement Directive. Adopted on 29 November 2017: p10.
28. DPA 2018. s10 (3) and Part 2 of Sch 1.
29. Proust O, Defromont J-A. Post-GDPR French Data Protection Law adopted. Fieldfisher Privacy, Security and Information Law Blog; 11 Sept 2019.
30. European Parliamentary Research Service. How the General Data Protection Regulation changes the rules for scientific research; July 2019.
31. Van Quathem K, Cooper D. European Commission Issues Updated Q&A on Interplay between the GDPR and the Clinical Trials Regulation. Covington & Burling Inside Privacy; 15 April 2019.
32. S.I. No. 314 of 2018 (Ireland) Data Protection Act 2018 (Section 36(2)) (Health Research) Regulations 2018; Regulations 3(1)(e) & 5(5).
33. DPA 2018. Sch 2, para 27, (2).
34. DPA 2018. Sch 2, para 27, (1).
35. Elteste U, Van Quathem K. German court decides on the scope of GDPR right of access. Covington & Burling Inside Privacy; 8th August 2019.
36. DPA 2018. Sch 2, para 16, (1).
37. DPA 2018. Sch 2, para 16, (2) (a).
38. DPA 2018. Sch 2, para 16, (2) (b).
39. This does not include another health professional; DPA 2018, Sch 2 Para 17 (2).
40. DPA 2018. Sch 2, para 16, (4) (b) (ii).
41. DPA 2018. Sch 2, para 27, (1) & (3).
42. The Information Commissioner's Office. Guide to the General Data Protection Regulation (GDPR). ICO.
43. DPA 2018. Sch 2, para 27 (1).
44. GDPR. Art 17 (3) (c).
45. GDPR. Art 17 (3) (d).
46. GDPR. Art 17 (1) (b).
47. GDPR. Art 17 (2).
48. Taylor MJ, Wallace SE, Prictor M. United Kingdom: transfers of genomic data to third countries. Human Genetics. 2018; 137(8): 637–645.
49. Hackle E. GDPR: Decision of the DPA on the erasure of personal data. Lexology. 14th Nov 2019.
50. Ausloos J, Mahieu R, Veale M. Getting Data Subject Rights Right. LawArXiv, 25th Nov 2019; pp19-21.
51. Article 29 Data Protection Working Party. Guidelines on the right to data portability (WP 242) Adopted 13th December 2016. p8.
52. Taylor MJ, Wallace SE, Prictor M. United Kingdom: transfers of genomic data to third countries. Human Genetics. 2018; 137(8): 637–645.
53. GDPR. Art 21(1).
54. Or if it is in the substantial public interest on the basis of Member State law (Art 22 (4)) and safeguards are in place.

55. European Data Protection Board. Opinion 22/2018 on the draft list of the competent supervisory authority of the United Kingdom regarding the processing operations subject to the requirement of a data protection impact assessment (Article 35.4 GDPR). Adopted on 25th September 2018.
56. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Adopted on 4 April 2017.
57. European Commission. Adequacy decisions: How the EU determines if a non-EU country has an adequate level of data protection. ec.europa.
58. Biobanking and BioMolecular resources Research Infrastructure-European Research Infrastucture Consortium. GDPR Code of Conduct.
59. European Data Protection Board. Guidelines 2/2018 on derogations of Article 49 under Regulation 2016/679. Adopted on 25 May 2018.
60. Clarke N, Vale G, Reeves EP, *et al*. GDPR: an impediment to research? Irish Journal of Medical Science. 2019; 9(1): 1129–1135.

phg foundation
making science work for health

PHG Foundation
2 Worts Causeway
Cambridge
CB1 8RN

+44 (0) 1223 761900

@phgfoundation

www.phgfoundation.org

UNIVERSITY OF CAMBRIDGE