phg foundation
making science work for health

The GDPR and genomic data

The impact of the GDPR and DPA 2018 on genomic healthcare and research

Executive summary

ico. Information Commissioner's Office
Grant Recipient

Authors

Colin Mitchell, Johan Ordish, Emma Johnson, Tanya Brigden and Alison Hall

Acknowledgements

The GDPR and genomic data project was funded by the Information Commissioner's Office (ICO). We thank the ICO for their support.

May 2020

URLs were correct as of April 2020

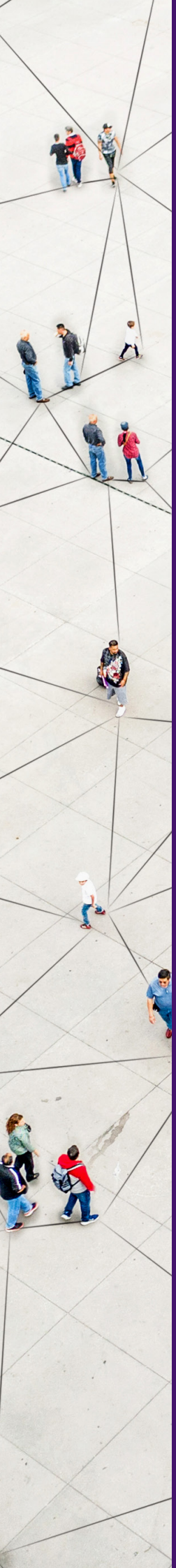
This report can be downloaded from www.phgfoundation.org

Published by PHG Foundation

intelligence@phgfoundation.org

UNIVERSITY OF CAMBRIDGE

Contents



Executive summary

Objectives and approach

The EU General Data Protection Regulation (GDPR) came into force in May 2018. It aims to strengthen and update the data protection framework across the EU and to respond to the challenges of novel technologies in the digital age. However, since revisions to the data protection regime were first suggested, the genomics community have raised concerns about their potential impact on the processing and sharing of genomic data.

With support from the Information Commissioner's Office (ICO), the PHG Foundation conducted research to assess the current and likely near future impact of the GDPR and UK Data Protection Act 2018 on uses of genetic/genomic data in healthcare and health research. Its objective was to identify and evaluate aspects of the GDPR and related UK legislation that are currently challenging for uses of genomic data in healthcare or health research, and to anticipate those which could soon become significant. This research also aimed to identify measures to mitigate or reduce negative impacts on genomic data processing, including opportunities presented by the GDPR for the development of appropriate standards in the genomics context.

Through legal analysis, key stakeholder interviews and a multidisciplinary policy workshop, this research identifies and evaluates a wide range of ways in which the GDPR and related Member State legislation impacts the processing of genomic data in healthcare or research.

In this report, key impacts are grouped according to significant parts of the Regulation (Chapters 3-7). Cross cutting issues and potential mitigations are discussed towards the end of the report in Chapter 8 and as part of the conclusions in Chapter 9.

When and where does the GDPR apply?

Two of the impacts raised in our research relate to uncertainty in determining precisely when the GDPR applies to genomic research. This is due to uncertainty about the breadth of the territorial scope of the GDPR and how the concept of a joint data controller should be interpreted in the context of genomics collaborations. The territorial scope (contained in Article 3) could apply to processing genomic data outside the EU/EEA. This might arise if an EU-based institution is part of a research collaboration processing the personal data of individuals from elsewhere in the world and if there is a sufficient link between that processing and the activities of the EU-institution. It is currently unclear what level of connection is required between processing outside the EU/EEA and the activities of the data controller within EU/EEA to bring such processing within the territorial scope of the GDPR.

The second uncertainty about the scope of data controllership is closely connected: if it is determined that an EU/EEA based institution or individual is a 'data controller' under the GDPR, it is possible that processing of genomic data outside the EU/EEA by another party will be within the scope of the GDPR. It is unclear when an individual or group has a sufficient involvement in determining the 'purposes and means' of processing personal data to constitute a data controller and this is potentially challenging for a wide range of actors in the genomics context. For example, committees advising on which data should be made available as part of genomics research and how that data should be shared, could be said to be influencing the purposes and means of processing that data. The Court of Justice of the European Union (CJEU) seems to have taken an expansive approach to data controllership in recent case law and it is unclear if this approach may be transposed to the genomics context.

Key point

The genomics and scientific research communities require more detailed guidance about territorial scope and the concept of (joint) data controllership taking into account examples from the genomics context. This will require collaboration with authorities such as the European Data Protection Board (EDPB) and national Supervisory Authorities (SAs).

When are genetic or genomic data 'personal data'?

The GDPR governs 'personal data', which means any information relating to an identified or identifiable natural person (Art 4(1)). Although this definition has not altered dramatically under the GDPR, our research identified uncertainty and disagreement whether genetic/genomic data and associated health data fall within the scope of the GDPR. In addition, the GDPR has expressly incorporated a category of 'genetic data' for the first time. However, there is further ambiguity about which data, resulting from what forms of analysis, fall within this definition. These uncertainties are already causing a substantial impact on genomics projects and challenging local, national and international flows of genomic and health data.

There are a number of associated challenges. One is in recognising when data that may not be assumed to be identifying becomes personal data due to technical advances, increasing availability of external data sources, for example, rapidly expanding public genealogy databases, or even through changes to the background knowledge of those who can access the data. Another challenge is that it is difficult to determine and agree when genomic data and associated clinical data have been sufficiently de-identified to fall outside the definition of 'personal data'. This is problematic in the genomics context where such data are frequently highly technical and multi-dimensional, and where there is some disagreement about the residual risk of identification following the application of different de-identification techniques (discussed in Chapter 8).

Further confusion has been caused by the inclusion of pseudonymised data (where data are coded and separated from additional identifying information) in the GPDR leading some to question whether pseudonymised data always remain personal data. We think there is no coherent basis for treating such data differently to other types of de-identified data. The same test of identifiability should be applied so that it is possible for pseudonymised data to become anonymised data if sufficiently safeguarded *in context*.

Other impacts relate to how genetic and genomic data are considered under the GDPR. One is the potential that some genomic data are treated as inherently identifying by the courts in England and Wales. We believe that this would be an incorrect interpretation of the GDPR and that a distinction between individuation (or singling out) and identification should be maintained. This requires consideration by SAs and the EDPB.

Key point

Determining when genetic or genomic data are 'personal data' will remain extremely challenging but this can only be assessed in context. The genomics community should be proactive in developing appropriate standards for de-identification of genomic data through a code of conduct or certification scheme setting out best practice for specific contexts and forms of data. This could help build consensus and achieve harmonisation of national and international approaches under the GDPR given the potential that such a code or certification scheme may be formally recognised under the GDPR (see further below).

Ensuring lawful processing of genomic data for healthcare and research

The GDPR requires a lawful basis for the processing of personal data (Article 6) and that additional conditions are met for processing special categories of data, such as genetic or health data (Article 9). Consent is one of the options to enable both types of processing but one of the significant impacts of the GDPR identified in the literature and in the course of our research, is the strengthened requirements for consent and the resultant challenges of relying on this lawful basis, in particular for genomics research.

One such challenge is that consent is presumed not to be freely given and is therefore not valid where there is a 'clear imbalance' between the data subject and the controller. Another is that consent should be specific and the scope for broader consent in the research context is limited.

Our analysis concludes that these are not necessarily insurmountable barriers to the use of consent and that there is scope at this early stage for the genomics community to argue for consent standards that are appropriate to genomics activities. Broader consent may be necessary and justified in certain areas of genomic research.

However, other legal bases and Art 9 conditions are available. In particular there are provisions for scientific research processing under Art 9. An important impact for cross-border genomics initiatives is that these may not be available in other Member States and may involve uncertain or inconsistent safeguards and additional requirements. In complex uses of genomic data that span healthcare, testing, technology development and research purposes, data controllers may need to identify multiple lawful bases and Art 9 conditions for different purposes.

One resultant challenge is that these bases and conditions have different consequences for data subject rights under the GDPR. The choice of legal basis for processing and Art 9 condition will make a significant difference to whether certain rights apply in a given context. Data controllers should carefully consider the available legal bases and consequent obligations for fulfilling data subject rights.

Key point

For data protection purposes, consent is not necessarily the most appropriate basis for processing personal genomic or health data. In the UK, clear alternatives are available for healthcare, public health and scientific research purposes but for cross-border collaborations, national variation means that the research community should continue to evaluate and advocate how GDPR requirements for consent apply in this context.

Fulfilling data subjects' rights and meeting obligations under the GDPR

Our research identifies a number of challenges and ambiguities in relation to the fulfilment of GDPR data subject rights in the genomics context. Because Member States are allowed to introduce exceptions and derogations to a number of rights (for example, in the case of scientific research), it is a challenge determining which rights apply in a specific context within and between Member States.

There are further challenges, such as interpreting and applying the exception for processing which does not require identification of a data subject (Art 11). High standards for reliance on this exception could, in fact, reduce efforts to de-identify or minimise data rather than promote increased efforts. Fulfilling the range of data subject rights under the GDPR requires significant effort. There are particular challenges in the genomics context. One is determining how to deal with the right to access to data (Art 15) which relate to several genetic relatives and how to manage and reconcile the interests of multiple family members in the same genomic information. Another is determining how the Art 16 right to rectification of inaccurate data applies to genomic data. One possibility is that the right to rectification could be applied to require records to be updated if their results are clearly no longer accurate, such as where there has been a change in status in the pathogenicity of a genetic variant.

Other challenges relating to data subject rights such as determining when research exemptions apply and whether the right to erasure may be fulfilled through anonymisation are also likely to arise in the genomics context.

Further obligations placed on data controllers and processors by the GDPR are significant in the genomics context. These include obligations to incorporate data protection by design, put in place safeguards to protect the privacy and security of data that are proportionate to the risks involved and to keep these under review as technologies and techniques progress. These safeguards encompass technical, organisational and legal measures and may require significant oversight and investment.

Key point

At this relatively early stage in substantiating the requirements of some data subject rights in this context, there is a valuable opportunity for the genomics community to proactively engage with SAs and the EDPB, to ensure that forthcoming guidance and standards governing data subject rights are relevant for the genomics context.

Challenges for genomic data sharing

Some of the major concerns about the effect of the GDPR relate to its impact on genomic data sharing. Since data sharing is a cornerstone of genomics, it is particularly impacted by the uncertainty and interpretative disagreements about data protection law described above. Brexit is set to further complicate and challenge data-sharing between the EU and the UK.

The genomics and research communities are already pursuing ways of streamlining and clarifying data sharing rules but policymakers, regulators and legislators need to be aware of the potential costs of the legal ambiguities identified in our research for healthcare and research. Parts of the GDPR have a powerful impact on the transfer of identifiable genomic and health data outside the EU/EEA and to international organisations, which is unlawful without a specific legal mechanism. Our analysis highlights that the selection of the appropriate legal mechanism is highly contingent upon the context and legal position of the controller.

There is a structured process for selecting a legal mechanism for international transfer. If an Art 45 adequacy decision exists (currently there are only 13 in place), this must be used before reaching for Article 46 safeguards or, failing those, Article 49 derogations. In terms of these safeguards and derogations, there is potential for the 'legally binding and enforceable instruments between public authorities' safeguard (Article 46(2)(a)) to be particularly useful for public authorities (such as university researchers) because it has recently been given a flexible interpretation by the European Data Protection Board (EDPB).

The safeguards on contractual clauses (Articles 46(2)(c) and (d)) also represent feasible solutions to transfer genomic data internationally, especially if a standard set of clauses could be secured for the sector as a whole. The safeguards on codes of conduct and certification mechanisms (Articles 46(2)(e) and (f)) also represent sector-wide methods to facilitate international data transfer. These mechanisms, although lacking precedent and difficult to secure, could have a key role in demonstrating general compliance with the GDPR for the sector as a whole, even if they are insufficient to act as a legal mechanism for international transfer. In exceptional circumstances, the Article 49 derogations may be useful.

The exception (or derogation) on consent (Article 49(1)(a)) will facilitate international data transfer under most conditions. However, consent for the processing of data under the GDPR in healthcare and research settings is set at a high bar and may be difficult to meet. The exception for important reasons of public interest (Article 49(1)(d)), while also set at a high bar, is a promising mechanism for the genomics community, given that the community will often be in a strong position to demonstrate this through the tangible and critical benefits from international cooperation. In all cases, an adequate level of protection 'essentially equivalent' to that guaranteed within the EU must be provided (Article 44). This means that any mechanism will be assessed according to how well it safeguards EU data subjects' fundamental rights in practice.

Key point

International collaborations underpin many areas of genetic and genomic research. In order to foster these collaborations, and support high quality, robust research, we recommend that the genomic sector should work toward sector-specific solutions such as standard contractual clauses, codes of conduct, and certification mechanisms to facilitate international transfer of genomic data.

The reduction and mitigation of potential deleterious impacts

Although this research has identified a range of potentially challenging impacts following the implementation of the GDPR we also identify some promising mitigations and measures that may address potentially deleterious impacts. For example, there is an active field developing strategies to mitigate or reduce the likelihood of re-identification of individuals while facilitating the processing of genomic data. This includes technical approaches, bringing analysis to data rather than sharing the data itself, using advanced computational methods such as homomorphic encryption to carry out secure analysis of sensitive data.

These approaches also combine technical measures with legal and organisational safeguards including data access controls, contracts, and internal policies, to help minimise the risks of misuse. However, the high level of protection required for genomic data and the technical nature of the risks and mitigations involved may be challenging for some controllers. For this reason, we conclude that there may be an increasing need for trusted third party processors who can process data applying technical and organisational best practice to generate the usable data that is needed for healthcare or research purposes.

Perhaps most promisingly, within the GDPR itself there is scope for the genomics community, or sub-sectors of genomics processors, to develop best practice and appropriate standards demonstrating compliance with the GDPR. These are codes of conduct (Article 40) or certification mechanisms (Article 42) which may be approved by SAs if they can be shown to meet a particular need of that sector or processing activity, facilitate the application of the GDPR, and provide sufficient safeguards and effective mechanisms for monitoring compliance.

These may not be easy or quick to develop, and the broader the coverage of codes or certification schemes, the more they will require an understanding of the variations in Member State law relating to processing of genetic and health data and the more difficult it may be to secure support for, and reach agreement on, their substance.

To try and reach consensus more quickly and to establish some level of certainty for genomic data processing under the GDPR, we suggest that the genomics sector could adopt a dual approach to codes of conduct: simultaneously pursuing a broad and sector wide code of conduct to establish and harmonise rules for genomic data as a self-regulatory code of conduct and more specific sub-sector-led codes that aim to crystallise best practice into smaller formal codes of conduct to be approved under Art 40.

Key point

Developing best practice for compliance with the GDPR requires on-going attention from genomic data processors and controllers. Both formal collaboration (codes of conduct or certification) and informal collaboration among the genomics community, health and research policymakers, data protection experts and supervisory authorities will be crucial to ensure that proportionate and appropriate standards are developed for the processing of genomic and health data.

Conclusions

Not all genetic information will be personal data or 'genetic data' under the GDPR and the Regulation does not create an exceptional regime of rules for genetic data which is distinct from other special categories of data that are subject to higher protection. Nevertheless, our research demonstrates that the GDPR impacts—or is likely to impact—genomic medicine and research in a very wide range of ways. A lack of legal certainty in a number of key areas challenges the ability of data controllers to comply with the GDPR and DPA 2018, and undermines potential collaborations. These fundamental challenges were the main focus of our workshop and interviews.

Our analysis also highlights potential impacts which are likely to create challenges in the near future. These include difficulties reconciling multiple data subject's rights over the same 'shared' genomic data and determining how specific principles and rights, such as data accuracy and the right to rectification, should be applied to genomic data.

There are further likely challenges for international data sharing following the end of the Brexit transition period, which will trigger the need for a legal mechanism to transfer personal data from the EU/EEA to the UK and vice versa. The potential for two forms of the GPDR (the EU and UK GDPR) governing processing in the UK and transfers between the UK and EU/EEA, will create increased scope for conflict and regulatory divergence.

For supervisory authorities, policymakers and the genomics community, addressing these impacts and providing guidance about the application of generally a sector-agnostic Regulation to the highly technical and specific context of genomics is a considerable challenge. This will require on-going, close collaboration between experts in genomic medicine, research and data protection at both national and international levels. The development of codes of conduct and certification could act as a form of co-regulation between SAs and specific sectors and provide the vehicle for addressing the challenges for genomic healthcare or research set out in this report.

Achieving consensus for key aspects of genomic data processing could improve harmonisation and legal certainty, not only for the genomics community but also for related areas, such as biomedical research.

The PHG Foundation is a non-profit think tank with a special focus on how genomics and other emerging health technologies can provide more effective, personalised healthcare and deliver improvements in health for patients and citizens.

For more information contact:
intelligence@phgfoundation.org

UNIVERSITY OF CAMBRIDGE

phg foundation
making science work for health