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

Annexe: policy workshop and interviews

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.

Published by PHG Foundation

intelligence@phgfoundation.org

The PHG Foundation is a health policy think-tank and linked exempt charity of the University of Cambridge. We work to achieve better health through the responsible and evidence based application of biomedical science.

We are a registered company, no. 5823194

UNIVERSITY OF CAMBRIDGE

Policy workshop and interviews

In order to supplement the analysis and evaluation performed as part of the GDPR and genomic data project, 13 interviews were undertaken with 15 interviewees. The interviewees were purposively selected to elaborate on a particular area of interest, or because they represented a particular stakeholder group. These interviews took place between July and December 2019. The interviewees and the findings from the interviews are described in section 2 of this annexe.

A policy workshop was also held as part of the project on 16th January 2020. In advance of the policy workshop, two discussion papers were drafted to provide background briefings to delegates and to provoke and support discussion. These papers, *Genomic Medicine and Research: when does the GDPR apply?* and *Genomic Medicine and Research: how does the GDPR apply?* are available separately.

The objectives of the policy workshop, the delegates, the proceedings and the key findings from the workshop are set out in section 1 of this annexe.

In addition, key comments from the interviews and the delegates are seeded throughout the GDPR and genomic data report where pertinent to the content. The comments from the policy workshop and from the interviews are not intended to be attributable to any specific delegate or interviewee, as this was one of the conditions of their participation in this research.

This document forms part of a larger project, GDPR and genomic data available at:

www.phgfoundation.org/research/data-protection-genomic-data

1 Policy workshop

A policy workshop entitled 'The impact of the GDPR on genomic medicine and research' was held on 16 January 2020 at the Franks & Steel Rooms, Wellcome Collection, 183 Euston Road, London NW1 2BE. The purpose of this workshop was to present the preliminary findings from the research, and to explore the impact of the GDPR on a broad group of stakeholders. The workshop started at 09.30 and ended at 16.00. As a consequence of the workshop being held under Chatham House Rules, this report addresses the key findings from the workshop but not their provenance. An agenda for the workshop is included below.

Invitees

Thirty-one delegates attended the workshop representing a broad cross-section of different stakeholder groups: clinicians, laboratory scientists, academics, researchers, legal professionals, representatives from patient groups and the commercial sector. In addition, five members of the PHG Foundation were involved in presenting, note-taking and organisation.

Materials provided to delegates in advance of the workshop

In advance of the workshop, delegates were sent the full agenda, joining instructions and two discussion papers. The discussion papers were drafted to provide background information to delegates to support structured discussion:

Discussion paper – Genomic Medicine and Research: when does the GDPR apply?

The EU's General Data Protection Regulation (GDPR) gives rise to significant uncertainty for those working in genomic medicine and research in Europe and those in collaborations involving EU citizens' data. In this discussion paper we address the challenge of understanding whether the GDPR applies to the data and the stakeholders involved in genomics projects. We outline one approach to this question and highlight the definitions of 'personal data', 'pseudonymisation' and (joint)'control' which will influence when uses of genetic and clinical data will be governed by the GDPR. If this is the case, researchers and health care professionals are required to support individuals' rights and fulfil the obligations set out in the GDPR.

Discussion paper – Genomic Medicine and Research: how does the GDPR apply?

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.

Proceedings

The introductory session confirmed that the objectives for the day were to understand the current and future impacts of the GDPR for genomic medicine and research; to identify what can be done to mitigate or reduce any negative impacts and to identify the priorities for clarification, that should be brought to the attention of policymakers and regulators.

Sessions

The GDPR and its impact on genomic medicine and research (Colin Mitchell)

The purpose of this session was to provide a general background to the GDPR, and the DPA 2018; the definitions of key terms within those provisions; and the current nature of genomic medicine in order to facilitate an understanding of the nature, scope and limits to 'genetic data' and to frame the questions and challenges to be addressed during the day.

Perspectives from practice

In order to frame the session and prompt discussion, some delegates were asked to elaborate on different aspects of the impact of the GDPR on their practice:

- Dr Helen Firth (Decipher) described how patient genotype and phenotype are combined in the DECIPHER genome browser and the processes for de-identified, coded and pseudonymised data. Key challenges are identifying when data are 'personal data' and ensuring a proportionate level of governance that balances appropriate data minimisation and protection whilst not compromising the utility of genetic and genomic data in facilitating patient diagnoses and enhanced patient management
- David Birkinshaw and Grant Stapleton (Genomics England) described the data flows utilised by Genomics England to deliver the 100,000 Genomes Project and the legal bases for processing these data. Focusing on data subject rights, they described how these were met within the project which straddles research and clinical care contexts.

Discussion and priority gathering

Over lunch, delegates were asked to rank the key challenges arising from the GDPR. These replicated many of the issues already identified and included:

- The scope of personal data in the context of genetics and genomics
- De- identification, pseudonymisation and anonymisation
- The scope of 'joint controllership'
- Interface issues including education and public engagement
- How to achieve greater harmonisation both nationally and internationally
- Proportionality and reducing the regulatory burden for research
- Interaction with the common law duty of confidentiality

Key challenges

Presentations by PHG Foundation staff focused on three key areas for practitioners and stakeholders, and were intended as a prompt for wider discussion.

Identifiability, pseudonymisation and anonymisation (Colin Mitchell)

This session addressed when genomic or genomic data are 'personal data'; when coded or pseudonymised data are 'personal data' and finally – how de-identification can be achieved and maintained.

Key discussion points:

- There is complexity and uncertainty about the scope of personal data
- Conflicting views as to whether pseudonymised data are always 'personal data' and the circumstances in which it might be regarded as more or less identifiable
- That the UK is able to adopt a more nuanced approach to de-identification for research because UK legislation facilitates processing of data provided that there are appropriate safeguards in place
- More work is needed to clarify how the reasonable likelihood standard should be interpreted for genomic medicine and research
- The difference between identification and individuation is complex: identification describes where a person could be singled out when data is combined with other data sources
- Controllers also need to consider the potential for re-identification through consideration of environment and context. Imputation is a strong tool in some contexts e.g. Iceland
- The more widely available the information, the wider the pool of 'determined' actors may be

Facilitating data subjects' rights (Colin Mitchell)

This presentation highlighted the breadth of relevant data subjects' rights, including the rights to information, access, portability, rectification and erasure. It went on to explore how these rights could be met in the context of genetic and genomic medicine and research and considered the potential limitations of Article 11.

Key discussion points:

- Does the right of access necessitate disclosure of a full copy of personal data?
- Variant classification and the need for rectification to comply with the right to 'accuracy' under the GDPR
- Interpretation of scientific research – what is the threshold for 'prevent or seriously impair'
- How useful is Article 11 for the genetics/genomics sector?
- Fulfilling Art 17 through anonymisation?

International data sharing (Johan Ordish)

This session explored the legal mechanisms available for data transfers to third countries or international organisations and discussed which might be most useful for the genomics community.

Key discussion points:

- There is a hierarchy of potential mechanisms ranging from an adequacy decision, appropriate safeguards, derogations and compelling legitimate interests. Many of these have conditions applied to their use
- Compelling legitimate interests may be available in the absence of other mechanisms but are limited in scale and scope
- Codes of conduct might be useful and the genomics community should be proactive in developing these. The 'cloud' code of conduct could provide an exemplar for understanding the practicalities involved. However, codes require a degree of consensus that might be difficult to achieve
- Sector specific contractual clauses may be more feasible, starting with provisions in the current standard contractual clauses which are seen as problematic

Further challenges

- Legal cases apply to specific contexts and their potential impact on the genomics sector could be distinguished on this basis
- Joint controllership does not necessarily mean equal responsibility: there may instead be a granularity of control

Ways forward for policy, regulation and practice

This concluding session was an opportunity for the delegates to highlight the priorities for future policy development.

Key findings:

- Clarification or guidance is required about many aspects, including the following:
  - Status of coded or pseudonymised data
  - Scope of (joint) 'data controller' in genomic collaborations
  - Use of consent as a legal basis for research
  - How practice might vary where research use includes commercial partners, especially regarding public expectations
- Codes of conduct or certification schemes might offer opportunities for legal harmonisation, and can demonstrate that appropriate safeguards for international transfers are complied with, or, with suitable endorsement, monitoring/compliance (Art 41)
  - Starting with a narrow scope can be helpful in building consensus and achieve quicker progress. Promising areas for discussion include principles for re-identification and transparency, which might help to promote public trust in the structures and processes involved

Key discussion points:

- Stakeholders should come together to respond to calls for input from statutory authorities and regulatory authorities. These collaborations should include research funders, international research organisations and patients/research participants
- Whilst achieving greater interpretative clarity might be at the expense of reducing flexibility, developing use cases within guidance might be helpful
- Developing more intelligent methods of data access that interrogate data and transfer results rather than necessitating data transfers might facilitate increasingly targeted and proportionate data flows
- Patient and participant expectations of how data are used are important but often missing from this debate
- 'Myths' about the chilling effect of the GDPR are a powerful deterrent to international data transfer and can be addressed by building on examples of best practice
- When developing sector specific approaches, we should guard against genetic exceptionalism

List of attendees

We thank the following attendees for their time, engagement, and generous sharing of ideas:

Name	Job title	Affiliation/Organisation
Shirlene Badger	Patient Advocacy Lead (EMEA)	Illumina
Regina Becker	Strategy Development Advisor ELIXIR-LU & Bioinformatics	Université du Luxembourg
David Birkinshaw	Data Protection Manager	Genomics England
Ruth Boardman	Partner, Privacy & Data Protection	Bird & Bird
Sarion Bowers	Head of Policy	Wellcome Sanger Institute
Laura Bradford	Senior Research Associate	Centre for Law, Medicine and Life Sciences, University of Cambridge
Phil Bradley-Schmieg	Associate	Bird & Bird
Vicky Chico	Lecturer in Law/Data Policy Advisor	Health Research Authority
Louise Coleman	Policy Analyst	Genetic Alliance UK
Paul Comerford	Principal Technology Advisor	Information Commissioner's Office

Name	Job title	Affiliation/Organisation
Sarah Dickson	Head of MRC Regulatory Support Centre	Medical Research Council
Edward Dove	Lecturer in Health Law and Regulation	University of Edinburgh
David Erdos	Senior Lecturer in Law and the Open Society, Faculty of Law	University of Cambridge
Helen Firth	Consultant Clinical Geneticist	Cambridge University Hospitals Trust
Julia Foreman	DECIPHER Project Manager	Wellcome Sanger Institute
Spencer Gibson	Research Associate	University of Leicester
Suzannah Gozna	Senior Legal Counsel	Wellcome Sanger Institute
James Groake	Associate Legal Counsel	Wellcome Trust
Daniel Harris	Policy Officer (Standards and Ethics)	General Medical Council
Naomi Hawkins	Associate Professor	University of Exeter Law School
Anneke Lucassen	Professor of Clinical Genetics Honorary Consultant in Clinical Genetics	University of Southampton Wessex Clinical Genetics Service
Tariq Malik	Lead - Office for Data Release	Public Health England
Nadia Meliti	General Counsel	Wellcome Sanger Institute
Robert McCombe	Senior Policy Officer	Information Commissioner's Office
Simon Ramsden	Chair Consultant Clinical Scientist	ACGS Manchester Centre for Genomic Medicine
Jonathan Sellors	Legal Counsel	UK Biobank
Grant Stapleton	Head of Service Delivery and Senior Information Risk Owner	Genomics England
Ross Thornton	Senior Project Manager	Office of the National Data Guardian
Rajoo Veeren	Senior Policy Advisor	Office of the National Data Guardian
Marc Wadsley	Research Associate	University of Leicester
Susan Wallace	Honorary Lecturer of Population and Public Health Sciences	University of Leicester

Agenda

The impact of the GDPR on genomic medicine and research: challenges and ways forwards

09:30	Registration and coffee	
10:00	**Welcome and overview**	**Alison Hall** PHG Foundation
10:15	**The GDPR and its impact on genomic medicine and research**	**Colin Mitchell** PHG Foundation
10:45	**Impacts: perspectives from practice**	**Helen Firth and Julia Foreman** DECIPHER
11:00	**Coffee break**	
11:15	**Impacts: perspectives from practice**	**David Birkinshaw and Grant Stapleton** Genomics England
11:30	**Discussion and priority gathering**	**Chaired by Alison Hall and Colin Mitchell**
12:15	**Lunch**	
13:00	**Key challenges** Identifiability, psuedonymisation and anonymisation Facilitating data subjects' rights	**Colin Mitchell** PHG Foundation
	International data sharing Further challenges	**Johan Ordish** PHG Foundation
15:00	**Break**	
15:15	**Ways forward for policy, regulation and practice**	**Alison Hall** PHG Foundation
16:00	**Close**	

2 Interviews

As part of the GDPR and genomic data research programme, 13 interviews were conducted with 15 interviewees.

We thank all the interviewees for their time and input into the GDPR and genomic data project.

Purpose of the interviews

Interviews were conducted for the following purposes

1. To add detail in relation to specific areas of legal interpretation or questions
2. To sense check preliminary findings
3. To develop a clear and representative understanding of the potential impacts on research and healthcare and to understand how these may be mitigated
4. To develop the recommendations and conclusions from GDPR and genomic data project, and understand how these might fit in the wider policy landscape

Key findings from the interviews are highlighted throughout the report

List of interviewees

We thank the following interviewees for their time, engagement and for sharing their ideas so generously:

Name	Job title	Affiliation/Organisation
Mark Bale	Head of Science Partnerships Deputy Director, Science Research and Evidence Directorate	Genomics England UK Department of Health and Social Care
David Birkinshaw	Data Protection Manager	Genomics England
Sarion Bowers	Head of Policy	Wellcome Sanger Institute
Yves-Alexandre de Montjoye	Assistant Professor, Director of the Computational Privacy Group	Imperial College, London
Robert Eiss	Senior Advisor to the NIH Director	Fogarty International Center, National Institute of Health, USA
Helen Firth	Consultant Clinical Geneticist	Cambridge University Hospitals Trust
Julian Foreman	DECIPHER Project Manager	Wellcome Sanger Institute
Suzannah Gozna	Senior Legal Counsel	Wellcome Sanger Institute
Shawn Liu	Director, Legal (Product and Privacy)	Chan Zuckerberg Initiative

Name	Job title	Affiliation/Organisation
Anneke Lucassen	Professor of Clinical Genetics	University of Southampton
Nadia Meliti	General Counsel	Wellcome Sanger Institute
Simon Ramsden	Chair Consultant Clinical scientist	Association of Clinical Genetic Science Manchester Centre for Genomic Medicine
Jonathan Sellors	Legal Counsel	UK Biobank
Grant Stapleton	Head of Service Delivery and Senior Information Risk Owner (SIRO)	Genomics England
Caroline Wright	Associate Professor in Human Genetics and Genomics	University of Exeter College of Medicine and Health

Key points from interviews

Current challenges

- Limitations and challenges in using the legal basis of consent.
- Framing privacy notices in ways that are transparent and accessible
- A lack of interpretative clarity in the GDPR which stems from divergent national cultures and frameworks but which will lead to regulatory divergence
- That this lack of clarity will cause some data processors to take a precautionary approach and leads to a 'ex-post facto application of the GDPR' which treats anonymised data as potential pseudonymised data in the light of big data analytics and new technologies
- Concerns about how research sponsors can meet data subjects' rights to withdraw consent and continue to meet other administrative, legal and ethical obligations (e.g. for adverse event reporting)
- For international genomic organisations that there will be a challenge in trying to ensure that the right safeguards are in place for data subjects from different Member States that meet MS interpretations
- That statistical methods aimed at protecting privacy such as releasing a small part of a dataset, may be inadequate to protect privacy
- Transparency is important: giving an overall average risk masks the potential for a subset of people to not have their privacy guaranteed to a minimum level.
- There are limitations to international data sharing, particularly around the use of cloud-based storage providers. These use standard agreements, but the impact on the sector is uncertain so far
- The characteristics of genomic data is that they are aggregated, dynamic and 'practically impossible to shift to another institution'
- There is considerable variability of data involved – including data relating to dead people (therefore not personal data); relating to people who have never lived (cell lines) and even synthetic data (as a way to extract meaning and obfuscate potential identifiers)

Future challenges

- How specific must consent for future uses be?
- Functionality that allows interpretation in real time rather than data uploads
- A concern that data subject access rights could be used at scale to access genomic data

Radical solutions

- The acknowledgement by regulators that broad consent is appropriate in some situations
- More systematic use of model licences
- Instead of giving access to individual level data, query-based systems can upload a code that can be run against the dataset allowing inferences to be made. This is suitable for extremely structured data. Synthetic datasets can also be used instead of releasing personal data
- Data use ontology may be machine generated and automated in the future
- Approaches which measure the privacy preserving characteristics of systems in terms of their resistance to adversarial attacks
- That genomic organisations limit their sphere of activity to selected MSs
- A sector specific approach that takes account of the challenges for the research community
- Rather than 'owning' genetic data, in a research context this should be viewed as a 'non-rivalrous' data commons – part of the common heritage of mankind

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