GDPR
What are the implications for health-related research?

A Farr Institute Brief

Overview

This brief has been produced as part of the work of the Farr Institute Information Governance Working Group. It offers a **Need-to-Know Guide** to the European Union's General Data Protection Regulation (GDPR) as this relates to the conduct of health-related research. This document draws on a larger report that considers the issues in full and which is available here.

Background

The EU General Data Protection Regulation came into force on 25 May 2018 and is directly applicable in all the member states, regulating a broad scope and range of data processing activities. For the UK, the Data Protection Act 2018 ensures GDPR application in the post-Brexit era. The Data Protection Act 2018 and the GDPR should be read together.

This brief examines the GDPR key changes for health-related research. It also advises on the main implications of the GDPR and how the research community should address them when processing both personal data and special categories of personal data for health-related purposes. While the GDPR is the latest legislative initiative to impact on health-related research, it is important that this is also read together with other legal measures, such as the common law duty of confidentiality.

We share the opinion that the GDPR is the outcome of reaching a fair balance between data subjects' rights and research community's interests, and that it provides for a stringent data protection framework. We also acknowledge that this might impair research if either undue reliance is placed on consent as the means to conduct data-driven research, or if the GDPR risk-based approach and compliance provisions are undermined. Nonetheless, we strongly believe that the GDPR should be seen as an opportunity rather than a threat to research. The GDPR represents a model of other routes to lawful research beyond consent, such as those in the public interest, and these are discussed herein.

Key pointers

- Research is broadly defined under the GDPR (rec. 159). The GDPR adopts a broad and inclusive definition of research, including both private, public and (non) commercial research activities by public and/or private entities. Public health is considered a sub-category of scientific research.
- Personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with the requirement of article 89(1) to apply the appropriate technical and organisational measures (art.5(1)(e)).

- Data controllers should be in a position to sufficiently identify and record the reasons that require and justify a longer storage period of personal data.
- The Information Commissioner's Office suggests that personal data could be stored and retained for an indefinite period of time, provided that they will be processed solely for research purposes and under the safeguard of the appropriate technical and organizational measures. In this case, this data processing should not result in any decisions affecting particular individuals.
- The distinction between anonymous and personal data remains. Thus, if anonymous data is processed, then the GDPR does not apply. It is therefore of crucial importance to understand what is meant by 'personal data' and 'anonymous data' under the GDPR.

The GDPR introduces the principle of accountability (art. 5(2) strengthening the obligations of data controllers who must be in a position to demonstrate compliance with the GDPR. Additionally, the GDPR introduces direct obligations and responsibilities for data processors (see further below).

Definitions

'Personal data' means any information relating to an identified or identifiable natural living person ('data subject') who can be directly or indirectly identified in particular by reference to an identifier, such as NHS numbers and names. Thus, pseudonymised data are likely also caught by this definition (see further below).

The GDPR also refers to 'special categories of data'. Special categories of data refer to data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation. Due to their sensitive character the GDPR provides for more stringent requirements regarding their processing.

Anonymous data is information that does not relate to an identified or identifiable natural living person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.

Anonymisation: take home message

Processes of anonymisation and pseudonymisation are not a precise science. They are techniques in risk-minimisation and are subject to review over time as data are linked to other data. Indeed, they are context-dependent processes aiming at risk-management and not risk-aversion. A constant and ongoing assessment of the properties of information and its link with other data and data subjects is necessary. Therefore, data controllers and processors should NOT read these as fixed categories.

Data are more likely to be treated as 'anonymous' if re-identification would require unreasonable time, effort or resources, taking into consideration the available technology at the time of the processing and technological developments.

In order to determine whether a natural person is identifiable, account should be taken of all the means *reasonably likely* to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly.

Pseudonymisation is 'on the spectrum' and is not of itself determinative relevance. Pseudonymised data can produce anonymous data for third parties, provided that pseudonymisation is irreversible and re-identification is impossible as far as third parties are concerned.

Data controllers should:

- Conduct risk-assessment exercises of their data systems, justifying any claims to use of 'anonymised' data;
- Conduct ad hoc and context-dependent assessments on a regular basis;
- Apply additional technical and organisational measures and safeguards to ensure that the key-code is securely and safely kept separately from the pseudonymised data.

Do I need consent to conduct research on health-related data?

The GDPR sets higher consent standards than under the previous legislation, providing for a freely given, specific, informed, unambiguous and affirmative indication of the wishes of data subjects. The GDPR consent requirements aim to empower data subjects in terms of power and control over their data. Therefore, opt-out systems, implied consent and pre-ticked opt-in boxes do not meet the GDPR requirements and such policies should be updated.

The GDPR recognises the value of research for society and the potential inapplicability and impracticability of seeking consent in a research context. Recital 33 could be seen as an implicit recognition of broad consent for research purposes. It allows for deviations from the specification and granularity of consent requirements in scientific research. Broad consent should be obtained in conformity with recognised ethical standards for research, including approval by a recognised ethics committee.

This means that data subjects can consent to certain areas of research or parts of a research project, but not generally to 'research'. Research purposes need not be fully specified, but they must at least be well-described. Comprehensive research plans, increased transparency and regular updates about potential amendments in the research project could further enable controllers to meet these requirements.

Important notice: when processing special categories of data – i.e. sensitive data, then stricter rules apply. For example, this raises challenges if relying on broad consent. Researchers should be very circumspect when triggering the application of this flexibility in order to legitimise the processing of special categories of data. In any case, consent to the processing of special categories of data should be explicit.

Whereas consent is often sufficient to meet the requirement for ethically-robust research and ensures the genuine expression of the wishes of data subjects to participate in research, we would advise data controllers to establish a different lawful ground and avoid reliance on consent as a means to comply with the GDPR. It is unlikely that consent will be the most appropriate and convenient lawful ground for storing and using data for research.

Indeed, consent is not recommended when data is collected by public authorities, including universities and the NHS, due to the presumption of the power imbalance between data controllers and subjects. On the contrary, the lawful grounds of the public task and legitimate interests should be considered as the most appropriate lawful grounds when processing personal data for research purposes. More specifically, public authorities, such as NHS organisations and universities, are entitled to process personal data if this processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in them (art.6(1)(e)). Individuals and private entities, including charities organisations, could conduct data-driven health-related research based on the lawful ground of legitimate interests (art. 6(1)(f)).

Although more stringent rules apply regarding the processing of special categories of data, data processing in the research context could be carried out if the processing relates to personal data which are manifestly made public by the data subject (art. 9(2)(e)), the processing is necessary for reasons of substantial public interest (art. 9(2)(g)), reasons of public interest in the area of public health (art. 9(1)(i)), or when the data processing is necessary for the provision of health or social care or treatment or the management of health or social care systems and services (art.9(2)(h)). The most appropriate and convenient lawful ground to establish in the present case refers to processing of special categories of data necessary for the archiving purposes in the public interest, scientific or historical research purposes or statistical purposes (art. 9(2)(j)).

Consent: take home message

The take-home message is that consent is only one among the available lawful grounds and it does not enjoy a special status. Researchers should use other grounds for data processing for research purposes to avoid the shortcomings of the legal and practical requirements of consent.

Can we conduct research in the public interest?

As analysed above, public authorities, including NHS organisations and universities, can use the 'public interest' clause as a lawful ground for data processing, including for research purposes. Private entities may also perform a task carried out in the public interest or in the exercise of official authority based on national provisions.

Data controllers should take into consideration the following:

- Data controllers should demonstrate and justify the public interest by reference to their incorporated or statutory purpose, as expressed in their constitution, charter and/or based on the public purpose/function vested within their authority.

- Overall public interest refers to the common welfare, rights and interests of the public, that are worth being recognised and protected. Recall, due protection of citizens' confidentiality is also a matter of public interest.

As far as the common law of confidentiality is concerned, the disclosure of confidential information is permitted if it is justified in the public interest. Thus, a number of factors must be taken into account in any consideration as to whether the research justifies a breach of confidentiality in the public interest:

- the nature of the information must be considered
- the use that will be made of it
- how many people will have access to it
- the security arrangements to protect further disclosure
- the advice of an independent expert advisor, such as a Caldicott Guardian, should be sought
- the potential for harm or distress to patients.

Under the GDPR, the public interest or official authority vested in the data controller should be laid down by the EU or UK law, which will specifically regulate the processing activities carried out for reasons of public interest, including public health.

There is current uncertainty about whether research falls explicitly under the concept of public interest under the Data Protection Act 2018. Indeed, the 2018 Act clarifies what public interest compromised in section 8, without referring to research. However, we do not consider that research is excluded from this concept. Indeed, the wording of section 8 involves an enumeration of public tasks/interests that is only indicative, non-exhaustive and open-ended. Therefore, researchers could invoke the public interest basis in accordance with the above remarks about how public tasks and functions can be identified.

In order to use this lawful ground, data controllers should conduct a balancing test. This balancing test should be recorded in line with the principle of accountability and carried out in conformity with the necessity and proportionality requirements.

More specifically, data controllers should:

- identify and specify the relevant task, function or power and identify its basis in common law or statute.
- demonstrate that there are no other reasonable and less intrusive means to achieve this purpose.
- the lawful ground of processing should be communicated to data subjects prior to any data processing.

This last point means that privacy statements should be updated and expressly include any legal basis satisfying the criteria for applying the lawful ground of public interest.

In order for data controllers to process special categories of data, they must show that the data processing is necessary for reasons of substantial public interest (art. 9(2)(g)), reasons of public interest in the area of public health (art. 9(1)(i)), the provision of health or social care or treatment or the management of health or social care systems and services (art.9(2)(h)) or necessary for the archiving purposes in the public interest, scientific or historical research purposes or statistical purposes (art. 9(2)(j)). Thus, the above remarks about public interest should be also taken into account in examining this ground. In addition to the above, specific measures should be laid down by national law to protect the fundamental rights and the personal data of data subjects.

The requirement of 'substantial' public interest necessitates an in-depth assessment, articulation and documentation of this interest. For example, setting electronic health file systems, where health data collected by healthcare professionals are shared with other healthcare providers for reasons of direct care, were justified on the reasons of substantial public interest under the previous Data Protection Directive.

Schedule 1 Part 1 paragraph 4 of the UK's Data Protection Act 2018 provides that the processing of special categories of data for research purposes is permitted if the data processing is necessary for the research purposes, carried out under the appropriate safeguards of article 89(1) of the GDPR, and is in the public interest.

What are the rules if data were originally collected for a different purpose?

Research is significantly based on the sharing and linkage of data from various sources. Data-repurposing is about further processing data, even if it is collected and held by the same data controller. This is also the case when public and private entities request access to data already held by other public authorities and private companies.

Data collection is the initial purpose of data processing. Any data processing subsequent to data collection constitutes further processing. Further processing is allowed only if it is compatible with these initial purposes. This requires a substantial and careful factual and legal assessment and a compatibility test.

At first sight, it seems that this could significantly impair research restricting data sharing and re-purposing. Nonetheless, the GDPR provides for a broad exception from purpose limitation in data processing for research purposes (art. 5(1)(b)). Indeed, the GDPR actually enables any data processing in this regard under the following cumulative requirements:

1. the data processing is carried out solely for further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, and

2. appropriate safeguards apply in accordance with article 89(1).

The Information Commissioner's Office has pointed out that "the GDPR specifically says that further processing for the following purposes should be considered to be compatible lawful processing operations: archiving purposes in the public interest; scientific research purposes; and statistical purposes. Even if the processing for a new purpose is lawful, you will also need to consider whether it is fair and transparent, and give individuals information about the new purpose."

Further processing of data for research: take home message

Further processing for archiving purposes in the public interest, scientific or historical research or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes for which it was collected.

Navigating the flexibilities of the GDPR for research

Subject to the adoption of the appropriate safeguards, member states can legislate for certain restrictions on the application of the GDPR where personal data are processed for scientific or historical research purposes and in so far as such rights are likely to render impossible or seriously impair the achievement of the research, and such derogations are necessary for the fulfilment of the research. This is provided by articles 89(1) and 89(2) GDPR.

In turn, section 19 of the Data Protection Act 2018 provides:

Processing for research purposes is in accordance with article 89(1) GDPR provided that:

- Processing of personal data is permitted when it is necessary for research purposes.
- This data processing is not likely to cause substantial damage or substantial distress to a data subject.
- This data processing is not carried out for the purposes of measures or decisions with respect to a particular data subject, unless the purposes for which the processing is necessary include the purposes of approved medical research.

In such cases, rights to access, rectify, restrict, erasure and object to processing no longer apply to research data, to the extent that complying would prevent or seriously impair the achievement of research. As far as the restriction of the right to access is concerned, such restrictions are permitted under the additional requirement that the results of the research are not made available in a form which identifies a data subject (Schedule 2, Part 6, para 27 of the 2018 Act).

Neither the GDPR nor the DPA 2018 explain exhaustively what amounts to "appropriate safeguards". Under the GDPR data controllers should adopt technical and organisational measures to ensure compliance with the data minimisation principle. Moreover, data should be anonymised if possible, otherwise pseudonymised, in order to prevent the re-identification of subjects. Those measures should be seen as the minimum requirement to meet the objective of article 89(1), namely to safeguard the rights and freedoms of data subjects. The GDPR does not further specify these safeguards because mere compliance with an exhaustive list of safeguards would be contrary to the risk-based approach that underpins

the Regulation. Similarly, the choice of the appropriate safeguards should be the result of a delicate, dynamic and sophisticated risk-assessment in accordance with the above restrictions.

What does robust risk assessment look like?

To understand the notion of the appropriate safeguards and design GDPR-compliant safeguard policies it is necessary to conduct thorough risk assessments. Risk assessments should include the evaluation of positive and negative consequences from data processing operations, be they present, future or remote, but in any case, there should be an account of which risks are reasonably foreseeable. These safeguards should be strong and apt to minimise or even eradicate the risks for material and/or immaterial harms to data subjects. In this context, possible emotional impacts, such as distress, embarrassment and irritation and the infringement of fundamental rights, such as the right to privacy and non-discrimination, should be particularly examined. Moreover, data controllers should aim to ensure enhanced transparency and data minimisation and the exclusion of the possibility that decisions will be taken with respect to individuals.

It is important to reiterate that data protection is a 'grey area,' where mere compliance with a tool-kit is not sufficient. On the contrary, data controllers should scrutinise their research protocols and privacy policies and apply risk-assessments. The contribution of researchers is vital to this end, since they are expected to cooperate with the data protection officer and explain what is their actual role and how they use the collected data. This will enable an effective risk-assessment before and during any data processing activity.

To this end, the considerations provided for by the GDPR data protection impact assessment (DPIA) should be also carried out with regard to the appropriate safeguards. A DPIA is a mechanism for identifying and mitigating present and prospective risks in data processing operations, focusing on the possible risks for the rights of data subjects.

Where a type of processing, in particular, using new technologies and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data.

Regarding data-driven research, article 35 of the GDPR requires the conduct of a DPIA in the case of data processing on a large scale of special categories of data and a systematic monitoring of a publicly accessible area on a large scale. In addition to this, by virtue of article 35(4) the Information Commissioner's Officer (ICO) requires data controllers to carry out a DPIA if biometric and genetic data is processed, data is matched or combined from different sources and new technologies are employed. The ICO has also explicitly referred to data processing for research purposes as a case where a DPIA is necessary.

In this context, the data protection officer (DPO) plays a key role in helping data controllers to ensure compliance with the GDPR and, more specifically, to conduct a DPIA. The DPO is a

different position to Caldicott Guardians, who are responsible for protecting the confidentiality of people's health and care information. A DPIA must describe the nature, scope, context and purposes of the processing in order to identify potential risks for data subjects and subsequently identify appropriate measures in accordance with article 89(1) as well.

Similarly, a risk-identifiability assessment should be also carried out. This should be robust enough to identify and specify the possible risks, assess their probability and seriousness of, and the possibility of re-identification before any processing takes place. There should be procedures to record and document this procedure. Arguments about lack of infrastructures, financial and human resources will not be accepted as an excuse.

Following the above analysis, we advise data controllers to review the following indicative measures and consider whether they are applicable in your case in a case-by-case study:

- Conduct internal audits to investigate and assess internal and external factors that may pose a risk. This requires the cooperation between different departments within an organisation, so that existing internal controls, IT systems, legal procedures and policies are regularly reviewed.
- Restricted physical access to IT infrastructures and databases. All persons involved in data processing operations should be trained and, if possible, accredited.
- Data processors should have appropriate skills and experience technical infrastructure and data protection and auditing policies.
- Implementation of professional secrecy obligations.
- Access to IT systems only to a limited and authorised number of researchers.
- There is a plethora of indicatively suggested disclosure control measures to be taken prior to data processing:
  - Pseudonymisation
  - Data encryption
  - Data aggregation
  - Anonymization
  - Functional separation. It refers to the requirement to apply technical and organisational measures to ensure that personal data processed for research purposes cannot be processed for non-research purposes.
  - Attribute separation
  - Record suppression
  - Character masking
  - Generalisation
  - Swapping
  - Data Perturbation

Data processors: A key player under the GDPR

A data processor is a natural or legal person that processes personal data on behalf of the data controller. Under the Data Protection Directive data processors were only contractually liable against data controllers for breach of their contractual obligations. Indeed, the data

controller bore most obligations and responsibilities as the one who determines the purposes and means of the processing of personal data.

The GDPR now provides for direct obligations and responsibilities for data processors. The Regulation introduces specific and independent obligations and responsibilities for data processors. Nonetheless, the GDPR does not go that far to equate data controllers to data processors in terms of obligations and responsibilities.

Overall, both data controllers and processors must comply with the data protection principles. Data processors are now responsible for acting in accordance with the instructions of data controllers (art. 29) and the GDPR provisions. Indeed, processing by a processor shall be governed by a contract or other legal act under Union or Member State law. Given that such written agreements should have already been concluded between data controllers and processors, you should update them or conclude new agreements meeting the GDPR requirements. The content and provisions of these agreements are provided in detail by the GDPR, including the subject-matter and duration of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects and the obligations and rights of the controller (art. 28(3)).

Moreover, under the GDPR data processors are responsible for implementing appropriate technical and organizational measures to secure personal data, such as pseudonymisation and encryption, (art. 28(1) and 32), data breach notification obligations (art. 33(2)), appointing an EU representative (art. 27) and a data protection officer (art. 37) under certain conditions and keeping records of data processing (art. 30(2)). There are also restrictions regarding data transfers to non-EU countries by data processors (art. 44-49).

The GDPR also provides for stringent restrictions to prevent data processors from engaging sub-processors without the prior specific or general written authorization of the data controller. In the case of a general written authorization to engage sub-processors, the data processor must inform the data controller of any changes to sub-processors and give the data controller the opportunity to object to these changes (art. 28(2)). In this case, the obligations of the initial data processors equally fall on sub-processors (art. 28(4)). Sub-processors must also comply with the instructions of the data controller (art. 32(4)). Therefore, data processors should act with due diligence when processing data and designating sub-processors, since the initial data processor remains liable to the data controller for the performance of the sub-processors' obligations (art. 28(4)).

In case of non-compliance with the GDPR data controllers and processors may be held directly liable and may suffer administrative fines, private lawsuits and criminal penalties (art.77-84).

Practical guidance – A short data protection toolkit

Data controllers should follow the below steps to ensure compliance with the GDPR:

- Appoint a Data Protection Officer. Data processors should designate one where necessary.
- Conduct data protection impact assessments when necessary.
- Existing privacy impact assessments must be reviewed, updated and aligned with the GDPR.
- Rule of thumb: Anonymise or apply the GPDR (keep anonymisation under review).
- Identify the necessary lawful grounds and applicable derogations, before any processing is carried out.
- Review lawful grounds for processing and if consent is needed, determine how this will be obtained and documented.
- Review and amend current data privacy notices and policies.
- Review, update and establish clear and GDPR-compliant data retention, management and deletion policies, documents and procedures.
- Consider whether data should be pseudonymised or whether it is necessary to apply additional or alternative security techniques and safeguards.
- Review and update processing agreements between controllers and processors.
- Review IT systems and policies to ensure data protection and privacy by design and data minimisation.
- Periodically check your IT system setting the confidentiality, integrity and safety of data as priorities.
- Develop efficient IT system, policies, and procedures to handle data subjects' requests.
- Update your IT system to prevent unauthorised access to data and cyber attacks.
- Develop and test a real-time data breach response mechanism.
- Provide compliance training to any personnel processing personal data.
- Do not apply a 'release and forget' privacy policy, but constantly update your policies and align them with the ICO guidance.
- Remember, it is not just all about the GDPR, and that other legislative and ethical requirements must also be met, such as the duty of confidentiality.

Adam Panagiotopoulos
School of Law
University of Edinburgh
August 2018