

Justice Select Committee: Inquiry on EU Data Protection Framework Proposals

Response by the Wellcome Trust

August 2012

KEY POINTS

- The Government must make the protection of research one of their priorities in negotiations on the Regulation.
- It is essential that Article 83 and associated derogations are maintained as the Regulation moves through the legislative process. Amendments to clarify and strengthen the research provisions would be beneficial to ensure these achieve their intended purpose and do not inhibit important health research.
- Amendments are needed to ensure that the use of pseudonymised data in health research is regulated proportionately and to ensure clarity in the scope of the Regulation.

INTRODUCTION

1. We welcome the opportunity to respond to this inquiry since it is vital that the EU and UK can establish a regulatory framework that balances the rights and interests of individuals with the societal benefits of research using patient information. Our response focuses on the aspects of the proposed Regulation that affect health research. We are also submitting a joint statement from the Trust and other health research organisations that was presented to the Ministry of Justice during their call for evidence. This statement sets out the impacts of the data protection proposals for the sector and includes a number of case studies.
2. Information from patient records provides the foundation for much health research, and offers significant potential to answer questions about the factors that influence health and disease. Information from patient records can be used for epidemiological research; to understand more about the causes of disease; to detect outbreaks of infectious diseases; to monitor the safety and efficacy of drugs and medical devices; and to study the effectiveness of treatments and interventions. Patient information is also used to identify participants for research studies. Researchers may wish to approach individuals in order to gain their consent to participating in a particular piece of research, for example the trial of a new treatment for a particular disease.

Will the proposed Regulation strike the right balance between the need, on the one hand, for a proportionate, practicable but effective system of data protection in the EU, and on the other for business and public authorities not to be stifled by regulatory, financial and administrative burdens placed upon them?

Research derogations

3. The Regulation provides a number of derogations from particular requirements for the use of 'personal data' for scientific research, providing that personal data is processed in accordance with the conditions set out in Article 83. These derogations do not exempt research studies from all the requirements set out in the Regulation. The Wellcome Trust warmly welcomes this approach since it provides a framework that balances the facilitation of research with the protection of the interests of research participants. However, to safeguard this balance the Government must prioritise the protection of Article 83 and ensure the associated derogations for research are protected as the Regulation moves through the legislative process.
4. There are a number of issues around Article 83 and the associated derogations that would benefit from clarification to better reflect the intent of the clauses. The lack of clarity in the current UK Data Protection Act has contributed to a risk-averse culture among those sharing and using data for research, which has led to delays to important research.
5. In order to avoid replicating these difficulties, it is essential that any lack of clarity is rectified in the new Regulation. The following clarifications are needed:
 - Clarification of Article 6.4 and Recital 40 to ensure that the processing of personal data for other purposes intends scientific research to be viewed as a compatible purpose in itself.
 - Clarification that the reference to Article 83 (processing for historical, statistical and scientific research purposes) within Article 81 (processing of personal data concerning health) is intended to link the two sections, rather than to impose an additional restriction on research.
6. A number of aspects of the research requirements and derogations rely on demonstrating 'necessity'.¹ While this approach is reasonable in principle, it will be important that an appropriate and consistent definition of 'necessity' can be applied in this context to ensure clarity and proportionality in implementation.

Scope of the Regulation

7. The scope of the Regulation is 'personal data' that identifies a natural person, or from which a natural person can be identified.² It is important that the research community is clear about when the different types of data used in research – anonymised data; key-coded or pseudonymised data; and identifiable data (see Annex A) – are considered to be "personal data". This determines whether a research study is brought within the remit of the Data Protection Act and therefore must comply with its requirements. Clarity in the scope is essential so that those sharing and using patient data in research are fully

¹ For example Articles 6.2; 9.29(i); 17.3(c); 83.1(a); and 83.2(c).

² Articles 3 and 4

aware of their responsibilities, but do not impose unnecessary additional requirements that will stifle research.

8. The Regulation is not explicit on whether pseudonymised data are intended to be included within its scope. Pseudonymised or key-coded data underpin a substantial amount of research, for example studies at the Wellcome Trust Sanger Institute and the UK Biobank research resource. In the UK, the Information Commissioner has published draft guidance³ to the effect that pseudonymised data can be considered anonymous – where identification does not take place, or where identification does take place and the data protection principles are not breached – and therefore falls outside the scope of the Data Protection Act. Inclusion of pseudonymised data within the scope of the Regulation would therefore dramatically increase the regulatory burden on research.
9. The use of pseudonymised data in health research is well-established and operates within a system designed to reduce the possibility of re-identification of participants. It is important that the use of pseudonymised data in research is handled within a proportionate regulatory framework that takes into account the actual likelihood of re-identification under current conditions, not just the technical possibility of re-identification. Conditions that will reduce the actual likelihood of re-identification could include the use of ‘safe havens’, such as England’s new Clinical Practice Research Datalink and comparable services in the devolved nations; contractual data sharing agreements; and professional standards for researchers that prohibit re-identification. In many instances the identifying code will not be held at the research site where the pseudonymised data are used in research, but at a hospital or by a safe haven. The Regulation should be amended to provide greater clarity on this issue for research, for example by noting that conditions could be established in a Member State that preclude re-identification, therefore ensuring that re-identification would not be considered “reasonably likely”. The UK Government must ensure that the proposed Regulation does not increase the regulatory burden of using pseudonymised data in research.
10. Anonymous data falls outside of the scope of the Regulation. However, the act of removing identifiers to ensure that data are no longer personal – anonymisation – could fall within the definition of processing (Article 4). This would mean that the process of anonymisation itself would have to comply with the requirements of the Regulation to be lawful. We suggest that the Regulation should be revised to expressly permit anonymisation, while prohibiting re-identification for data that has been anonymised.
11. Clarification is needed around ‘genetic data’ and ‘data concerning health’ to ensure that these definitions are only intended to apply to personal data that falls within these categories, rather than all related data. Further, the definition of ‘data concerning health’ should be clarified and must be consistent with Recital 26 to make it clear that data concerning health does not include biological samples *per se*, but rather to personal data obtained from testing such material.

³ http://www.ico.gov.uk/about_us/consultations/our_consultations.aspx

Are the next steps the UK Government proposes to take during the negotiations, set out in the Summary of responses to its Call for evidence, the right approach?

12. The Government's *Summary of Responses* to the Call for Evidence recognises the issues for research in the draft Regulation (pp31-32). However, research is not reflected as a priority in the Government's proposed next steps. It is important that this is rectified to ensure that the draft Regulation does not hinder research in the public interest. Particular steps the UK Government must take to protect the balance between the rights and interests of individuals and the societal benefits of research using patient information, include:

- Protecting Article 83 and the associated derogations for research as the Regulation moves through the legislative process.
- Seeking amendments to clarify and strengthen the research provisions to ensure these achieve their intended purpose and do not inhibit important health research.
- Ensuring that the proposed Regulation does not increase the regulatory burden of using pseudonymised data in research.

The Wellcome Trust is a global charitable foundation dedicated to achieving extraordinary improvements in human and animal health. We support the brightest minds in biomedical research and the medical humanities. Our breadth of support includes public engagement, education and the application of research to improve health. We are independent of both political and commercial interests

ANNEX A: THE TYPES OF PATIENT DATA USED IN HEALTH RESEARCH

Health data can be accessed by researchers in the following forms:

- **Identifiable data** – these include information in patient records such as patients' names, addresses, dates of birth and NHS numbers. There are also aspects of health data that could become identifying when they relate to a diagnosis of a rare condition or when combined with other data. Identifiable data are needed when future contact is needed with the participant, for example to contact them to take part in a study, or to link information across different data sets.
- **Key-coded or pseudonymised data** – these cannot directly identify an individual, but are provided with an identifier that enables the patient's identity to be re-connected to the data by reference to a separate database containing the identifiers and identifiable data. Pseudonymised data can often be used in place of identifiable data.
- **Anonymised data** – these data cannot be connected to the original patient record. Anonymised data are suitable when no contact is needed with the participant or where the data does not need to be linked to any other data sources.