

Safeguards framework for secondary use of clinical trial data for scientific research

Author: Brendan Barnes



- Purpose of this paper:** Clinical data, including data generated in clinical trials, are valuable resources for scientific health research. Such data have the potential to yield therapeutic insights and discoveries, independently of whether they are used for the initial purpose of the trial or re-used later on for different health-related scientific research purposes. This paper is concerned with the re-use of clinical trial data for “secondary health research”, meaning scientific research performed for health-related purposes other than the drug development program or other objective for which the clinical trial was performed. In the GDPR, re-use of data is referred to as “further processing”. Under the GDPR, the re-use of clinical trial data for secondary scientific research purposes is subject to implementing appropriate safeguards to protect the rights and freedoms of data subjects. The objective of this EFPIA Data Governance Working Group paper is to describe the safeguards which are used when clinical trial data are re-used for secondary health research purposes in the pharmaceutical industry and by responsible organisations and researchers in other health sectors. The conclusion of this paper is that there are well-established safeguards in place. This paper does not discuss anonymisation of clinical trial data, considering that anonymised data do not fall within the scope of the GDPR.
- Relevance of safeguards to secondary scientific research under the GDPR:** The GDPR facilitates the use of personal data for scientific research by providing several exemptions to the usual rules applying to the processing of personal data. These exemptions recognise the potential societal benefits of using personal data for scientific research.
 - Under the GDPR, personal data may not be processed in a manner which is incompatible with the purposes for which they were collected. Processing personal data for scientific research purposes is not considered incompatible if appropriate safeguards are in place to protect data subjects’ rights and freedoms. This is an exemption to the “purpose limitation” principle for scientific research performed with appropriate safeguards;
 - The exception to the prohibition on processing health data and other “special categories” of personal data which Article 9.2(j) of the GDPR provides for scientific research purposes is also subject to the implementation of appropriate safeguards for data subjects; and
 - Similarly, the exception to the obligations to provide data subjects with detailed information about processing of personal data for scientific research purposes is subject to the implementation of safeguards.



The relevant provisions of the GDPR are:

*Article 5.1(b): "...further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, **in accordance with Article 89(1)**, not be considered to be incompatible with the initial purposes ('purpose limitation')."*

*Article 9: [Processing special categories of personal data is prohibited unless] 9.2(j) "processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes **in accordance with Article 89(1)** based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject."*

*Article 14.5(b): [The obligations to provide detailed information to data subjects about the processing of their personal data shall not apply where and insofar as] "the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, **subject to the conditions and safeguards referred to in Article 89(1)** [...] In such cases the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available."*

*Article 89.1: "Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, **shall be subject to appropriate safeguards**, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner."*

Each of the exemptions is subject to the implementation of safeguards to protect data subjects from the possible risks of the use of their personal data. Decisions regarding the safeguards adopted when using clinical trial data for secondary health research are therefore highly relevant to the question of whether the exemptions apply and therefore to the ability to realise the potential of re-using such data.

3. **Examples of secondary health research with clinical trial data:** The secondary use of personal health data generated in clinical trials is valuable for a broad range of scientific research purposes, as anticipated by Recital 159 of the GDPR: "...processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research. In addition, it should take into account the Union's objective under Article 179(1) TFEU of achieving a European Research Area. Scientific research purposes should also include studies conducted in the public interest in the area of public health..." Examples of health research purposes for which clinical trial data may be re-used are:

- 3.1 To test new hypotheses which could otherwise require a new research study involving medical intervention on patients. Such new hypotheses may for example be aimed at studying other therapies or improving understanding of disease mechanisms;
- 3.2 To improve the efficiency, design and methods of future clinical trials;
- 3.3 To enable validation or scrutiny of the original results by independent researchers;



EFPIA Brussels Office
Leopold Plaza Building * Rue du Trône 108
B-1050 Brussels * Belgium
Tel: + 32 (0)2 626 25 55 *
www.efpia.eu * info@efpia.eu



- 3.4 For regulatory purposes, such as to perform safety studies requested by authorities;
- 3.5 For research (often performed in collaborations) involving the pooling of clinical trial data from multiple studies to facilitate treatment evaluation and drug development; and
- 3.6 To develop and validate new health care technologies, such as algorithms based on artificial intelligence.

4. **Context of secondary health research with clinical trial data:** Secondary use of clinical trial data for health research takes place in a range of contexts, utilising different forms of personal data and safeguards. The research may be done:

- by the original sponsor/controller;
- within scientific collaborations in which the original sponsor/controller may or may not be a participant; or
- by third party organisations or researchers.

The data may be used in isolation or combined with other data sets. Enabling these different possibilities maximises the potential societal benefit of secondary use of clinical data, by facilitating collaborative efforts, fresh and creative approaches, and diverse technologies and disciplines to be applied to the data to unlock potential to yield new insights and therapies. The safeguards which are applied must be appropriate to the needs of the research in each case and must always protect the rights and interests of the data subjects.

5. **Safeguards when using clinical trial data for secondary health research:** Responsible companies, institutions and researchers use well-recognised safeguards to protect the rights and freedoms of clinical trial participants when they re-use clinical trial data for secondary health research. The safeguards must be seen in the context that the purpose of scientific research is to test hypotheses, not to reach conclusions about individuals or to make decisions affecting them personally such as their access to health or life insurance.

The following safeguards are applied when using clinical trial data for secondary health research:

- 5.1 Oversight of research assures the scientific merit and ethical integrity of research proposals. Such oversight may be conducted internally or externally, such as by ethics committees;
- 5.2 The clinical trial data used in the secondary research does not contain the clinical trial participants' direct identifiers such as their names or social security numbers. Such identifiers are excluded by the clinical trial site before transferring the data set to the sponsor. The participants' names are replaced by random codes assigned by the principal investigator pursuant to the requirement for key-coding under the ICH Good Clinical Practice Guideline E6(R1) 10 June 1996 (Sections 1.58 and 5.5.5). The key linking the code to the trial participant's name is held by the Principal Investigator under strict obligations of professional secrecy and is only accessible by other parties under very limited and regulated circumstances, such as by study monitors and by health authorities conducting inspections to evaluate the integrity of the clinical trial data. Key-coding is synonymous with pseudonymisation, which Article 89.1 specifically mentions as an example of an appropriate safeguard to protect data subjects. In addition to replacing the clinical trial participant's name with a code and removing other direct identifiers, the clinical trial site excludes indirect identifiers such as the address and contact details from the clinical trial data which are shared with the sponsor;
- 5.3 Appropriate security measures are applied when transferring, using and storing clinical trial data to:
 - prevent accidental or unlawful destruction, loss or alteration of the data;
 - prevent unauthorised access to the data or unauthorised disclosure of the data;
 - assure the integrity of the data; and
 - allow data access and use to be logged and monitored.



EFPIA Brussels Office
 Leopold Plaza Building * Rue du Trône 108
 B-1050 Brussels * Belgium
 Tel: + 32 (0)2 626 25 55 *
www.efpia.eu * info@efpia.eu



A clinical trial sponsor is subject to obligations under ICH Good Clinical Practice Guideline E6(R1) 10 June 1996 in relation to the electronic handling of clinical trial data. Logically, the obligations addressing data security must continue to apply when the data set is used for secondary health research. There is no justification for lowering the standard of protection for trial participants' privacy when the data are used for other purposes.

- 5.4 Due diligence is performed on collaboration partners and third party service providers to ensure that they have the necessary expertise and infrastructure to comply with data privacy requirements;
- 5.5 In data sharing agreements and research collaboration agreements, parties receiving or having access to the data for secondary research purposes are bound by contractual obligations including prohibitions on attempts to re-identify data subjects, confidentiality and non-use obligations, data security requirements, research purpose limitations, and reporting and documentation requirements;
- 5.6 The scope of the data used for research is minimised, so that only the data required for the purpose of the scientific research are used, and the data are only processed to the extent required for that purpose;
- 5.7 There are data protection programmes in place including policies, standard operating procedures, training, and compliance monitoring and audits. Through such programmes, organisations structure and formalise the knowledge and expertise, control mechanisms and accountabilities required to protect clinical trial participants' rights and freedoms when their clinical trial data are used for secondary health research. These procedures help to ensure that clinical trial data are not used or shared for direct marketing or other incompatible purposes;
- 5.8 There is oversight of data privacy compliance by data protection officers or other persons with similar responsibilities; and
- 5.9 Where health care professionals perform the research, there are professional obligations of secrecy towards clinical trial participants.

Organisations performing health research use Data Protection Impact Assessments where required by law, or other similar risk assessments, to fine tune the implementation of safeguards to address risks of the research to the clinical trial participants, taking into account all the relevant factors. These assessments may be specific to each research project, a group of comparable processing activities or the performance of health research using clinical trial data generally. The way safeguards are implemented for a particular research project may depend, for example, on the data subject population (e.g. children), the type of data used in the research (e.g. rare diseases), the data required to fulfill the scientific research purpose, as well as the context within which the data are shared and with whom they are shared (e.g. with a well-known, reputable research institution or university vs. multiple parties with little experience of working with health data). Other relevant factors which form part of the assessment include possible motivations to re-identify clinical trial participants and the time and cost which would be required to do so. In some cases, depending on the specificities of the scientific research project and the context in which it is performed, the Data Protection Impact Assessment may lead to further safeguards being applied when using clinical trial data for secondary health research:

- Secure data platforms are used which allow data to be interrogated and analysed without being physically transferred or copied; and
- Further de-identification of the key-coded clinical trial data, as an additional measure to protect clinical trial participants from being identified. Such measures may involve for example aggregating the data, introducing noise, or replacing specific values with ranges, while retaining the statistical value of the data to the extent possible so that they can be used to fulfill the research purpose.

6. **Safeguards are well established:** Responsible companies, institutions and researchers should only engage in scientific research using personal health data if they have the necessary infrastructure, resources and expertise to apply the necessary safeguards. The technical and organisational measures required to implement the safeguards described in this paper are well



EFPIA Brussels Office
Leopold Plaza Building * Rue du Trône 108
B-1050 Brussels * Belgium
Tel: + 32 (0)2 626 25 55 *
www.efpia.eu * info@efpia.eu



established in health research, such as in the performance of clinical trials, due to the high degree of regulation and scrutiny of the health care industry as well as auditing by regulators. The importance of protecting clinical trial participants and other patients is firmly embedded in the culture of organisations active in this field, as reflected in their standard processes and procedures for pharmacovigilance, medical information inquiry handling, as well as clinical trials and other health-related research activities.

7. **Data protection principles:** The use of clinical trial data for secondary scientific research performed using the safeguards described above upholds the data protection principles:
 - 7.1 **Fairness and lawfulness of data processing** is assured by protecting trial participants from being identified and thereby from being singled out, embarrassed, unfairly treated or discriminated against. By its nature, scientific research is not aimed at making decisions which affect individual data subjects. Ethical oversight is a further check against unfair discrimination, as well as to ensure that the clinical trial data are only used for appropriate scientific research needs;
 - 7.2 Using clinical trial data in a key-coded form (i.e. the only form in which the data set is stored by the sponsor) is an important aspect of **data minimisation**. Data minimisation is emphasised by Article 89.1 as a safeguard for protecting the rights and freedoms of data subjects, with pseudonymisation (synonymous with key-coding) given as a specific example of data minimisation;
 - 7.3 The documentation provided to clinical trial participants at the point of recruitment plays a key role in meeting **transparency** obligations and makes it clear that data may be used for secondary purposes, the legal basis under which this will take place and the possible relevant safeguards, including measures for de-identification. More detailed transparency about secondary research can be provided by making information about the secondary use of clinical trial data available on a public forum. In addition, companies publish much of their scientific research, thereby contributing to the public interest in scientific progress. Leading scientific journals typically require a data-sharing plan as a condition of publication. Clinical research legislation provides for the registration of studies and the publication of results, expert assessments and lay summaries. There is therefore a high level of societal transparency and accountability;
 - 7.4 Re-using personal data for purposes other than those for which they were collected may prove difficult or impossible under the GDPR because of the **purpose limitation** principle, but there is an exemption to this principle when personal data are used for scientific research. Article 5.1(b) adapts the purpose limitation principle when the scientific research is performed using appropriate safeguards. Re-using data under such circumstances is deemed to be compatible with the initial purpose of the data collection. Individual rights to protection of personal data are thereby balanced with societal interests, as reflected in recital 4 GDPR;
 - 7.5 Organisations carrying out clinical trials are required to retain data for regulatory purposes i.e. at least 25 years after the end of the clinical trial under the EU Clinical Trial Regulation No. 536/2014. The GDPR provides for retention of clinical trial data solely for secondary health research for longer than the legal retention period applying to clinical trial data (Article 5.1(e)). Doing so will be consistent with the **storage limitation** principle as long as safeguards are in place;
 - 7.6 The safeguards which are implemented assure appropriate security when transferring, using and storing clinical trial data for secondary health research, thereby upholding the privacy principle of **integrity and confidentiality**.
8. **Conclusion:** Re-using clinical trial data for drug discovery and other health-related scientific research facilitates the discovery and development of new and improved therapies, benefitting the lives of patients and, ultimately, society as a whole. The GDPR enables the use of clinical trial data for secondary scientific research purposes which are based on Union or Member State law, provided that appropriate safeguards are implemented to protect the rights and freedoms of the



EFPIA Brussels Office
Leopold Plaza Building * Rue du Trône 108
B-1050 Brussels * Belgium
Tel: + 32 (0)2 626 25 55 *
www.efpia.eu * info@efpia.eu



clinical trial participants. As described in this paper, responsible companies, institutions and researchers performing scientific health research implement well-established safeguards as part of their standard processes and procedures, tailoring the safeguards as necessary according to the risk-based approach mandated by the GDPR. In this context, the GDPR is a favourable framework for the re-use of clinical trial data for secondary scientific health research.



EFPIA Brussels Office
Leopold Plaza Building * Rue du Trône 108
B-1050 Brussels * Belgium
Tel: + 32 (0)2 626 25 55 *
www.efpia.eu * info@efpia.eu

