Public Consultation

Guidelines 1/2026 on processing of personal data for scientific research purposes

Part 2

Technical safeguards, pseudonymisation and digital infrastructures in health research

1 Contributors

The present contribution to the Public Consultation is jointly submitted by:

- Prof. Cláudia Monge (claudia.monge@fd.ulisboa.pt)
  Faculty of Law, University of Lisbon
  Lisbon, Portugal
- Prof Gunnar Piho, Associate Professor Medical Informatics (gunnar.piho@taltech.ee)
  Olga Vovk, Ph.D. student (olga.vovk@taltech.ee)
  EMedLab at TalTech, Department of Software Science
  Tallinn University of Technology, Estonia
- João Massano, MD (joao.massano@ulssjoao.min-saude.pt)
  Consultant neurologist, Head of Research and Clinical Trials
  ULS S. João – National Health Service Local Trust
  Porto, Portugal
- Vitor Ramos, MD (vramos@ensp@unl.pt)
  Lisbon, Portugal
- Joaquim Cunha (jcunha@healthportugal.com)
  Executive Director, Health Cluster Portugal (health sector association)
  Leça do Bailio, Portugal
- Eduardo Freire Rodrigues, MD (eduardo@uphillhealth.com)
  CEO Uphill Ltd (primary care software developer)
  Lisbon, Portugal
- Peter Villax (peter@villax.net)
  CEO Mediceus Ltd (health data software developer)
  Lisbon, Portugal

2 Introductory note

This is part 2 of Contributors' response with respect to the public consultation concerning these Guidelines, and it addresses technical safeguards. Part 1, submitted simultaneously, covers central concepts in the GDPR, and Part 3 contains the clinician's perspective.

GDPR requires the provision of adequate technical and organisational measures by the controller to ensure compliance with the law but does not give any guidance as to form or content, allowing the creation of measures which follow the state of the art, and in fact allow for considerable freedom to develop the most appropriate safeguards.

This flexibility introduced by the GDPR for the processing of personal data for scientific research purposes and reinforced by the present Guidelines is inseparable from the obligation to implement appropriate safeguards under Article 89(1) of the GDPR. The present contribution fully supports this principle and should be interpreted as supporting a governance model in which broader possibilities for further processing, data re-use and broad consent are compensated by robust technical, organisational and ethical safeguards. In the context of scientific research involving personal data, particularly health data, these safeguards may include:

- Technical measures: pseudonymisation and secure key management, data minimisation at extraction and access stages; secure processing environments and controlled remote access; privacy-enhancing technologies.
- Organisational measures: independent ethical review and DPO oversight; role-based access controls and auditability; transparency towards data subjects throughout the lifecycle of processing; mechanisms for objection, revocation, and withdrawal where applicable; governance procedures for controlled re-identification where clinically justified.

The evolution of large-scale health research infrastructures, including those supported by the European Health Data Space Regulation, requires a governance approach in which safeguards, proportionality and accountability play a central role in compensating for the practical limitations of strict purpose specification in long-term scientific research ecosystems.

Accordingly, this contribution supports an interpretation of the GDPR in which the assessment of compatibility, broad consent and further processing should increasingly consider the structural equivalence of safeguards and governance mechanisms protecting data subjects.

Appropriate, adaptive technical safeguards which can help the creation of that structural equivalence are covered in this Part 2.

3 Requests for clarification and improvement

3.1 On the value of pseudonymisation

EHDS advocates the preferential use of anonymisation by requiring the use of pseudonymisation to be justified by controllers. Pseudonymisation may be the method of choice when following patients through life.

3.1.1 Issues

Anonymisation, which breaks the link between data and data subject in an irreversible manner, may prevent important information from being transmitted to its owner, in the situation described in section 3.2.1 below. When used as a technical means to create a privacy-

safe data lake, the utility of the database progressively diminishes, as there is no possibility to add new data to the records of anonymous data subjects. Anonymisation prevents longitudinal research (following each patient individually) that is updated with new data, which can no longer be linked to its irreversibly anonymised owner. In the field of health, it also prevents tracing back a patient to transmit new vital information.

In addition, the Guidelines should also recognise that, in longitudinal health research, pseudonymisation may be necessary not only to protect the identity of research participants, but also to preserve the patient pathway over time. This requires the ability to link newly generated data to the same protected clinical trajectory, without revealing the direct identity of the data subject, provided that the residual risk of re-identification is assessed, documented and reduced to an acceptable level through appropriate technical and organisational safeguards.

3.1.2 Proposal

When considering the benefits of data re-use in scientific research studies, the Guidelines should recognise the value of pseudonymisation as the proportionate default where re-contact may be clinically necessary and authorise:

a) the long-term storage and processing of pseudonymised personal health data where a need may arise in the future to re-associate it with their owners;
b) the means to make health data that is obtained for primary use, also available for secondary use in a way which allows new information obtained in scientific research to be correctly associated and shared with the data subject whose personal data was used in the research, and who need to be informed of any vital discovery.

Such measures would enable the extraction of valuable information from the longitudinal correlation of data and the use of this information for preventive and precision medicine and health policies. Moreover, pseudonymisation also helps to preserve the patient pathway over time, linking newly generated or acquired data to the same protected clinical trajectory.

A new paragraph is proposed:

158b. *Where controllers use identified health data stored in a primary use data infrastructure for individual health care purposes, and the same infrastructure is used to provide data subsets for use in scientific research studies, the data shall be identified by a pseudonym to allow research findings which may contribute to better care and better decisions by the health care professional treating a specific patient, to be retransmitted to the primary use data infrastructure.*

3.2 Protection vs. benefit in re-identification

The Guidelines are focused on protecting the data user against unwanted re-identification, without identifying situations where re-identification is useful or even vital. This may lead to the perception that re-identification is mostly detrimental.

3.2.1 Issues

The re-identification of data, even when deidentified, minimised and pseudonymised, and its association with the data subject's name may be beneficial or harmful, depending on the context of the re-identification. While the Guidelines stress the risks of re-identification (paragraphs 152, 161, 163, 170), they do not give commensurate importance to its benefits, for instance when scientific research discovers "a significant finding related to the health of a natural person", which needs to be communicated to the person or health professional treating that person, as stated in article 38(3) of the EHDS Regulation.

Re-identification may therefore be warranted in the case of clinical need. It should be managed with a specified governance procedure, appropriate key management and only occur where the data subject has expressed a wish to be informed of relevant clinical results or outcomes.

3.2.2 Proposal

The Guidelines should include a new paragraph to recognise the situations where re-identification is warranted and desirable.

> *161b. Where there is a reasonable expectation that a scientific research study will produce a significant finding related to the health of a person, which needs to be communicated to the attending doctor and to the data subject who has consented to re-identification and has indicated a wish to receive it, the controller should preferably use pseudonymisation to protect the identity of research study participants, and adopt adequate technical and organisational means to safely re-identify the data subject in case of need.*

3.3 Protection vs benefit with respect to mobile applications and digital tools

The Guidelines address means to share data of interest by persons less familiar with the use of mobile applications and digital tools.

3.3.1 Issues

There is ample support for the view that persons who are less comfortable with the use of digital devices need easier, more analogue means to access information, and thus not be discriminated in information access.

In this respect, paragraph 48 in the Guidelines suggests that the controller should disseminate information of interest to data subjects via a webpage or a newsletter (presumably an e-mail). However, emphasising these means may not encourage the development of other better suited technical solutions which address the needs of both the general population, as well as vulnerable or digitally challenged groups. Moreover, paper-based methods may not be the most suitable for consent management, and particularly less so for consent revocation.

Nevertheless, digital devices need to remain optional, and controllers should develop digital solutions which can also accommodate data subjects who prefer not to use them.

3.3.2 Proposal

If a controller operates a health data infrastructure which includes a digital device-based mobile application as means to interact and share data with the data subject, two concerns should be addressed.

Consent by electronic means. Consent for data processing may advantageously be collected by electronic means, even for data subjects who prefer not to use a digital device, insofar as it enables consent withdrawal to be effectively and quickly propagated by data controllers. In this case, controllers should offer adequate and easy means to digitally challenged data users to give consent by electronic means, by providing a person to assist in the consent process with a computer or a tablet at the point of registration by the data subject in a data infrastructure, in place of a mobile app. The controller's device could generate the data subject's confidential identifier pseudonym at the same time as consent for data processing is obtained.

Right to object. In the absence of a mobile app, communication with the data subject within the scope of scientific research and further processing may be affected, particularly in requesting a consent for research or recording an objection to further processing. Alternative means to contact eligible participants for scientific studies, such as a newsletter, or a referral by the data subject's doctor, may be costly and less effective than sending a message to the data subject's mobile app.

In the case of further processing, DPOs and regulators may consider an assessment of structural equivalence as described in Part 1, to account for the fact that it may not be possible for controllers to record the right of the data subject to object to further research.

Controllers should develop digital solutions which adapt protective measures to enable the participation of digitally challenged participants in scientific studies, and specifically the elderly and the vulnerable most likely to benefit from it, so that there is neither digital exclusion nor research exclusion.

A new paragraph is proposed:

> 56b. A c*ontroller should develop non-digital or assisted digital and organisational solutions which adapt protective measures to enable the participation, objection or consent withdrawal of digitally challenged participants in scientific studies, and specifically the elderly and the vulnerable most likely to benefit from it. The ethics committee or other appropriate authority may consider the quality of those solutions to assess the validity of further processing.*

3.4 On data becoming sensitive when associated with health data

New types of personal non-medical data will be increasingly used in scientific studies in association with health data. They may require the same type of privacy protection.

3.4.1 Issues

Personal non-medical data such as social media, fitness, nutrition, lifestyle, socio-demographic information, residence area and geolocation data may be used in association with health data to enrich the domain of study variables in scientific research and produce new knowledge and deeper insights into life and health outcomes. However these categories of personal non-medical data are not covered by article 9(1) of the GDPR, and, when associated with health data in scientific research, could have a similar impact on the privacy of the data subject. This issue is not addressed in the Guidelines.

3.4.2 Proposal

The Guidelines should recognise the value as well as the privacy impact of associating personal non-medical data with health data in scientific research. A new paragraph could be added to the Guidelines stating:

> xx. *When personal non-medical data, such as, for instance social media, fitness, nutrition, lifestyle, socio-demographic information, residence area or geolocation data are associated with personal health data for scientific research, they should be afforded safeguards equivalent to those applied to the special categories of personal data of Article 9(1) of the GDPR*.

4 Conclusion

The Contributors support the Guidelines' emphasis on safeguards under Article 89(1) of the GDPR and consider such safeguards to be central to the development of trustworthy scientific research infrastructures in Europe.

This contribution proposes that the Guidelines more explicitly recognise the operational and ethical value of technical safeguards such as pseudonymisation, secure processing environments, controlled re-identification procedures and modern digital consent-management tools. In particular, the Contributors consider that pseudonymisation may, in many health research environments, provide a more proportionate balance between scientific utility and privacy protection than irreversible anonymisation.

The contribution also highlights the increasing importance of contextual sensitivity, where non-medical personal data associated with health data may produce privacy risks equivalent to those of special categories of personal data under Article 9 of the GDPR.

The Contributors suggest that the Guidelines should encourage technically robust and ethically governed digital infrastructures capable of supporting longitudinal research, participant transparency, clinically relevant re-contact and secure secondary use of health data within the framework established by GDPR and the European Health Data Space Regulation.

Peter Villax, coordinator, Mediceus Ltd
peter@villax.net
23 June 2026