Response to the Public Consultation on the EDPB Guidelines 1/2026 on Processing of Personal Data for Scientific Research Purposes

We highly welcome the European Data Protection Board's initiative to develop dedicated guidelines on processing personal data for scientific research, as they hold the potential to clarify long-standing ambiguities surrounding broad consent, transparency, and data subject rights under Article 89(1) of the GDPR. Writing from an academic perspective, our comments aim to highlight the immense diversity across research fields—such as the profound methodological and structural differences between the humanities and the natural sciences—as well as the vast financial and technical disparities among institutions. Crucially, the final document must clearly distinguish between mandatory legal obligations and recommended best practices to prevent disproportionate administrative burdens, duplicate assessments, and unintended restrictions on the freedom of science.

While we fully understand and respect the document's primary focus on protecting the data privacy of data subjects, the current draft is significantly decoupled from the realities of everyday scientific practice. By evaluating research through a strictly theoretical and administrative lens, the guidelines largely overlook how modern science operates, particularly concerning the principles of Open Science, the technical limitations of researchers interacting with third-party platforms, and the well-documented burden of consent fatigue on research cohorts. To prevent these guidelines from becoming an insurmountable compliance barrier that stifles European innovation, it is vital to balance data protection with the operational needs of the scientific community. We strongly encourage the EDPB to revise the text to better integrate the perspective of scientific practice—specifically by providing pragmatic mapping frameworks for legal bases, addressing the necessity of data preservation in immutable repositories for scientific reproducibility, and anchoring requirements in realistic contextual analyses rather than abstract auditing duties.

1 Introduction

Current text in the Guidelines – section 3.2*: "Controllers can store personal data for scientific research purposes for longer periods of time, even if the original purposes for processing the data have been fulfilled. For example, if the initial research results give rise to future research projects that require additional processing of personal data, it may be permitted to store the personal data for such projects."*

While we fully recognize that setting a strict, universal retention period is neither feasible nor desirable—as certain types of invaluable, immutable data (such as genomic data) often require indefinite storage for long-term scientific utility—the undefined phrase "longer period of time" may cause significant misunderstandings.

Without further guidance, this ambiguity could create a sense of unease among data subjects who remain unaware of the expected duration of the processing.

Therefore, we recommend that the EDPB, rather than prescribing rigid time limits, provides clear criteria or examples of how controllers can transparently communicate estimated retention periods or the specific milestones that justify continued storage to data subjects.

Current text in the Guidelines - section 4: *"The GDPR does not exhaustively harmonise the requirements for the processing of personal data for scientific research purposes. Accordingly, the EDPB underlines that there is ample room left to, and consequently a need for, complementary Union and Member State (MS) law..."*

Comment: While we fully acknowledge and respect the legal architecture of the GDPR, which leaves room for national derogations under Article 89, we believe the EDPB's decision to completely leave these laws unaddressed in the guidelines is a missed opportunity. While legislative competence belongs to the Member States, providing non-binding, harmonized recommendations or a "best-practice framework" from the EDPB is highly desirable, if not essential, for several critical reasons:

1. The Internationalization and Transnational Nature of Modern Science

Scientific research is fundamentally borderless. Contemporary breakthroughs rarely happen within a single Member State; they rely on large, pan-European consortia, cross-border biobanks, and shared data spaces (such as the European Health Data Space). If every Member State constructs its own disparate legal framework for scientific data processing, it creates significant legal fragmentation. Harmonized guidelines from the EDPB would serve as a vital "common denominator," facilitating smoother cross-border data flows and preventing the administrative stagnation of international research projects.

2. The Disproportionate Burden on Smaller Member States

Smaller Member States often lack the institutional capacity, specialized legal resources, or historical precedents to draft highly sophisticated national laws for complex scientific data processing. By abstaining from providing recommendations, the EDPB inadvertently leaves smaller states to navigate these complexities in isolation. This can lead to either overly restrictive national laws (out of fear of non-compliance) or inadequate frameworks, both of which stifle local innovation and disadvantage researchers in smaller markets.

3. Legal Certainty for the Scientific Community

Researchers and scientific institutions are experts in science, not data protection law. The current "patchwork" approach forces academic institutions to conduct exhaustive

comparative legal analyses before launching collaborative projects. Clear, foundational recommendations from the EDPB would provide national legislators with a blueprint, leading to more predictable, cohesive, and user-friendly national laws across the EU.

Conclusion & Recommendation: We encourage the EDPB to reconsider this approach and to address these legislative gaps within its future guidance.

2 The scope of the concept 'processing of personal data for scientific research purposes'

2.1 Key-indicative factors for determining whether processing of personal data is motivated by scientific research purposes

Section 11 point VI. *"Potential to contribute to existing scientific knowledge or apply existing knowledge in novel ways. The research activities are merited, as they have the potential to contribute to existing scientific knowledge or apply existing knowledge in novel ways..."*

Comment: The current definition is insufficient as it remains overly broad and vague. Specifically, it fails to clearly define what constitutes "scientific potential" or "novel ways" of application. Given the rapid and unpredictable evolution of modern science, it is crucial that the guidelines do not attempt to establish a rigid or static legal definition of what qualifies as "novel." Instead, the final determination of scientific value and methodological innovation should ultimately be left to the scientific community itself, utilizing established peer-review mechanisms and expert evaluation. The role of the guidelines should be to provide a supportive framework, while respecting the autonomy of scientists to define the boundaries of their respective fields.

2.2 Research data infrastructures

Current text in the Guidelines: "*Personal data may be processed in a research data infrastructure (also referred to as a research repository or database)*"

We highly welcome the explicit inclusion of Research Data Infrastructures (RDIs) within the Guidelines, as recognizing these entities is crucial for modern, data-driven science. However, to avoid legal and operational ambiguities, we believe it would be highly beneficial to further clarify what structures and platforms are encompassed by the term RDI. To achieve this without burdening the text, we recommend that the EDPB proposes the creation and maintenance of an external, non-exhaustive registry or dynamic list of recognized RDIs. Keeping this repository separate from the main text would provide the necessary clarity while allowing the definition to be easily updated and adapted over time as research technologies evolve.

Current text in the Guidelines: "*If a researcher wants to access special categories of personal data for a research project, such as data on the health or mental well-being of*

pupils – pursuant to Article 9(2) of the GDPR – then it also needs ethical approval, in line with factor ii (adherence to ethical standards).”

To make this guidance truly effective, we would highly welcome a better and more comprehensive explanation of how ethical approval processes should properly function in practice. Specifically, it would be beneficial to clarify the operational interplay and the distinct roles of the individual researcher, the institutional ethics committees, and the research data infrastructures. Providing a broader framework or best-practice workflow for these approvals would significantly reduce ambiguity for all stakeholders involved.

3 Data protection principles (Article 5 GDPR)

3.1 Purpose limitation (Article 5(1) GDPR)

There is a potential for confusion regarding the boundaries between "further processing for scientific research purposes" and "ancillary processes for scientific research purposes." We recognize that rigidly codifying these terms could inadvertently restrict the flexibility needed for scientific advancement and evolving research methodologies. Instead of imposing rigid definitions, we recommend that the EDPB provides a set of practical, illustrative examples within the text. These examples should clearly delineate the practical boundaries between core scientific processing and purely ancillary administrative or technical support processes, thereby ensuring both legal predictability and the necessary room for scientific innovation.

3.2 Storage limitation (Article 5(1)(e) GDPR)

We highly welcome the recognition of data retention for verification and reproducibility purposes, and we propose that the Guidelines should provide clear criteria to justify such storage across diverse research fields.

4 Lawfulness (Articles 5(1)(a), 6(1) and 9 of the GDPR)

4.1 Consent (Article 6(1)(a) of the GDPR)

We highly appreciate the EDPB's guidance on broad and dynamic consent; however, the current text presents two major operational challenges that do not fully align with modern research practices. First, the guidelines largely overlook the practical reality of "consent fatigue." While continuous re-engagement under dynamic consent sounds ideal in theory, in practice, data subjects frequently become unresponsive or burdened by repeated requests. Forcing researchers to constantly re-seek granular consent for closely related secondary research risks high drop-out rates, which biases research cohorts and jeopardizes long-term scientific endeavours. We suggest placing greater emphasis on robust initial broad consent coupled with strong structural safeguards (such as independent oversight bodies) as a more viable solution.

Second, the text completely omits the critical intersection of data subject rights (e.g., withdrawal of consent) with data stored in immutable, open-access research

repositories (e.g., Zenodo). Modern scientific integrity, backed by funding agencies and journals, strictly mandates that exact datasets must be permanently preserved unchanged to ensure scientific reproducibility and peer verification. If a data subject withdraws consent after a study is published and the dataset is locked in an archive, a strict requirement for erasure creates an unresolvable conflict, as altering the archived dataset destroys its integrity and renders the research non-reproducible.

Therefore, we strongly recommend that the EDPB includes a dedicated sub-section addressing this friction point. The guidelines should provide pragmatic solutions for immutable repositories, explicitly clarifying how the principle of scientific reproducibility can be balanced with data subject rights under Article 89(2) of the GDPR, particularly by acknowledging that complete erasure may be restricted if it renders the verification of published scientific achievements impossible or seriously impaired. As a practical way forward, the guidelines could explicitly highlight the completion of a Data Protection Impact Assessment (DPIA) as an exemplary mechanism for controllers to formally assess, balance, and document these competing interests.

4.2 Public interest or exercise of official authority (Article 6(1)(e) GDPR)

Comment to par. 56.: We highly recommend that the Guidelines more explicitly instruct controllers to identify and understand the specific Member State legislation that further defines "public interest" within their respective jurisdictions. Furthermore, to ensure legal certainty, it is essential for the EDPB to clarify which legal basis applies, or remains binding, in scenarios where a Member State has not laid down a clear or specific legal framework for scientific research.

To improve the structural clarity and logical flow of the document, we also suggest moving the contextual information currently detailed in paragraph 59 to an earlier section, preferably aligning it directly with the core principles discussed in paragraph 56. This repositioning would significantly assist researchers and legal advisors in navigating the interplay between EU and national legislation.

4.3 Legitimate interest (Article 6(1)(f) GDPR)

Current text in the Guidelines -par. 61.: "*Accordingly, a controller that intends to process personal data for scientific research purposes on the basis of legitimate interest can, when applying the balancing test pursuant to Article 6(1)(f) GDPR, often attribute significant weight to the processing of personal data for scientific research purposes, in relation to the data subjects' interests or fundamental rights or freedoms.*"

While we welcome the clarification that scientific research can carry significant weight within the Article 6(1)(f) balancing test, we recommend that the EDPB explicitly addresses the interplay between this union-level legal basis and Member State law. Specifically, it should be clarified that while the legitimate interest assessment itself is

independent of national legislation, controllers are still fully bound by any specific national safeguards, conditions, or derogations enacted under Article 89 of the GDPR within the relevant Member State. Explicitly highlighting this boundary would prevent the misconception that relying on legitimate interest exempts a research project from mandatory national compliance requirements.

4.4 Processing special categories of personal data for scientific research purposes

Comment on Section 4.4 paragraph 65:

DPIA is mentioned in the whole document several times and more info is later in 8.2 chapter. Could you recommend specific examples of well-done DPIA as inspiration to the researchers reading the guideline? It can be only one with reference to documents where to find more information – maybe ARTICLE29 - Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01)

Comment on Section 4.4.2 (Article 9(2)(e) GDPR):

While we support the EDPB's goal of protecting data subjects on social media, the current requirements place an unreasonable and disproportionate burden on researchers. Expecting a scientific controller to verify the technical architecture of a platform—such as checking whether users can only post after adjusting specific settings with full knowledge of publicity (paragraph 70)—is operationally unrealistic. Researchers do not have access to the backend systems of social media platforms to audit user-interface flows, nor can they realistically evaluate the subjective state of mind or technical literacy of individual posters.

To make these guidelines useful and legally sound, the text should be anchored more precisely in established European case law. Instead of inventing abstract technical audit duties for researchers, the guidelines should directly rely on the specific criteria set out in C-252/21 (Meta v Bundeskartellamt) and C-446/21 (Schrems v Meta). The Court of Justice of the EU (CJEU) already provides a balanced framework: it looks at explicit, clear affirmative actions and the context of the publication, rather than forcing the researcher to act as an auditor of third-party platform designs.

Therefore, we recommend amending this section to shift the focus from platform technical verification to contextual analysis based on CJEU jurisprudence. The guidelines should clarify that a researcher satisfies Article 9(2)(e) GDPR if they verify that the content was, on the balance of probabilities, published by the data subject themselves on a profile or channel that is objectively open to the general public, without requiring an exhaustive investigation into the platform's onboarding settings.

Comment on Section 4.4.3 paragraph 75:

To help researchers better orient in all the footnotes and recommended documents it would help to write titles of the documents or paragraphs/articles or explain a bit more

what is written in the references. For example, in this paragraph: "One example of a provision in Union law that provides for a derogation pursuant to Article 9(2) GDPR is Article 53(1)(e) in the European Health Data Space Regulation (EHDS)" it could be more researcher-friendly: "*In this context, Article 53(1)(e) EHDS recognises scientific research as a valid derogation under Article 9(2) GDPR when processing health data for purposes such as innovation, development, and the training, testing, and evaluation of algorithms and digital health technologies, provided this contributes to public health, healthcare quality, or safety.*"

5 Obligations to inform (Articles 12-14 GDPR)

5.1 General

This article discusses the options available to data subjects regarding the granting and management of their consent to the processing of their personal data for scientific research purposes over longer periods of time.

Comment on section 5.1:

Data subjects should be able to easily manage their consent to the processing of their personal data. A solution that provides data subjects with easy-to-use channels to monitor what is happening with their personal data is fully compliant with the GDPR and helps build data subjects' trust and certainty.

In cases where multiple controllers are involved in scientific research activities, the creation of websites or applications should be coordinated.

The introduction of a privacy panel for data controllers, and where consent serves as the legal basis, the provision to the data subject of confirmation of consent, i.e., a digital record of the consent statement, would make it easier for data controllers to comply with the law. The data protection dashboard would enable data subjects to apply their rights under the GDPR.

5.2 Provision of information when personal data are collected directly from the data subject (Article 13(1) GDPR)

This article provides examples on how information should best be provided to data subjects in some situations where personal data is collected for scientific research purposes.

Comment on section 5.2:

This article provides several examples of how to properly inform data subjects in cases where whose personal data is directly collected from the data subject through a clinical trial, whose personal data is directly collected from the data subject through interviews, whose personal data is directly collected from the data subject through an application and when a controller does not have any contact details or direct contact with data subjects. The proposed solutions comply with current legislation (GDPR).

5.3 Further processing of personal data for further scientific purposes (Article 13(3) GDPR)

This article addresses situations in which a controller who has collected personal data for a purpose other than scientific research later decides to further process that personal data for the purposes of scientific research and the data subjects must be duly informed of this.

5.3 Further processing of personal data for further scientific purposes (Article 13(3) GDPR)
Comment on Section 5.4.2 (Article 14(5) GDPR):

While Section 5.4.2 outlines the standard exceptions to the information obligation under Article 14(5) GDPR, it presents a significant legal omission by failing to incorporate the critical jurisprudence of Case C-413/23 P. By omitting the principles established in this judgment, the current draft deprives scientific controllers of the updated legal framework necessary to navigate the complex threshold of when information exemptions can be validly invoked, particularly within broader institutional or regulatory procedures. Therefore, we strongly recommend that the EDPB explicitly integrates the baseline from Case C-413/23 P into this section, ensuring the guidelines provide an accurate, robust, and up-to-date guidance that reflects the current state of EU case law.

Comment on section 5.3:

If the controller has the ability to contact the data subject and subsequently obtain their consent for further processing of personal data for scientific research, it must do so.

A problem arises if public registers or other suitable means are not available, if their use would require a disproportionate effort, or if the data subject is not listed in any of the available sources; in such cases, the controller should inform data subjects indirectly.

When processing personal data for a purpose other than that for which it was originally collected, the controller must assess the compatibility of the original and new purposes in accordance with the GDPR. The data subject retains all of their rights at this stage.

I would recommend that data controllers consider whether a potential violation of the GDPR resulting from a failure to obtain consent from the data subject outweighs the interest in using the personal data for further scientific research.

5.4 Provision of information when personal data are not collected directly from the data subject (Article 14(1) GDPR)
Comment on Section 5.4.2 (Article 14(5) GDPR):

While Section 5.4.2 outlines the standard exceptions to the information obligation under Article 14(5) GDPR, it presents a significant legal omission by failing to

incorporate the critical jurisprudence of Case C-413/23 P. By omitting the principles established in this judgment, the current draft deprives scientific controllers of the updated legal framework necessary to navigate the complex threshold of when information exemptions can be validly invoked, particularly within broader institutional or regulatory procedures. Therefore, we strongly recommend that the EDPB explicitly integrates the baseline from Case C-413/23 P into this section, ensuring the guidelines provide an accurate, robust, and up-to-date guidance that reflects the current state of EU case law.

5.5 Changes to processing operations for scientific research purposes (Article 13(4) and 4(5)(a) GDPR) which data subjects have to be informed of

We welcome the lists of triggering and non-triggering changes in paragraphs 108 and 110, but the section does not address a scenario central to research data infrastructures: the onward provision of pseudonymised data as standing repositories admit successive accessing projects over time.

Paragraph 108 treats engaging new partners who "will receive personal data" as a triggering change. We recommend aligning this with Case C-413/23 P (EDPS v Single Resolution Board, 4 September 2025), which the draft does not cite. **The Court confirmed that pseudonymised data are not personal data "in all cases and for every person" (para 87), turning instead on the means reasonably likely to be used for re-identification.** Where a repository discloses genuinely pseudonymised data to a recipient who cannot re-identify the data subjects, those data may not be personal in that recipient's hands; conversely, the disclosing controller's transparency obligation remains fixed at the point of collection (paras 111–112).

This distinction determines whether admitting each new project is a "change" under paragraph 108 or falls within an already-communicated category of recipients under paragraph 110. We therefore recommend a worked example confirming that, where Article 89(1) safeguards and access governance ensure recipients cannot re-identify data subjects, such admissions may fall within paragraph 110.

6 Data subjects’ rights (Articles 15-21 GDPR)

Comment to par. 111: under Article 17(3)(c) regarding the inability to erase primary medical records for public health reasons, and Article 17(3)(d) for scientific research and statistical purposes, there are certain scenarios under Article 71 of Regulation (EU) 2025/327, that a patient's right to object under Article 21(6) GDPR becomes impossible to exercise, regardless of, and specifically under certain definitions of the purpose of utilisation (under section 3.1, specifically 3.1.2 of this document). This also renders informing the data subject pursuant to section 5.5, p. 107–109, and section 6.2.2, p. 121 of this document an informational obligation, not a safeguard according to paragraph 2 of section 1 referencing GDPR Article 89(1) and Article 8 CFREU. Specifics below:

6.1 General

Comment to par. 112: Data controllers can structurally force the invocation of Article 89(2) GDPR derogations. A purposefully designed ETL architecture for secondary use of health data for the purpose of research under paragraph 118 can potentially be used to legally invoke the rejection of data subjects' rights to erasure. A preventive provision should be implemented to discourage designed-not-to-comply as a default (p.119).

6.2 Right to erasure (right to be forgotten) (Article 17 GDPR)

Comment to par. 113: The guidelines do not clarify the exact timeline or "cut-off point" for when the exception under Article 17(3)(d) begins to apply in automated data processing environments.

In automated ETL (Extract, Transform, Load) pipelines, data is often extracted immediately after a research permit or project is approved. If the threshold of "seriously impairing the research" (para 120) is interpreted as applying from the exact moment of data extraction or permit approval, the data subject is effectively deprived of their right to erasure/object. They cannot object before the data is generated/extracted, and they are legally blocked immediately after.

Recommendation: The EDPB should clarify that the application of Article 17(3)(d) requires a balanced approach in automated systems. We propose introducing a mandatory "cool-down period" between data generation/availability and its actual extraction or integration into the research dataset. During this window, data subjects must have a practical opportunity to exercise their rights (e.g., via a patient portal).

This cool-down period could be waived only under specific circumstances, such as:

1. Where the data subject has provided explicit dynamic or broad consent.
2. In cases of overriding public interest, such as epidemics or public health emergencies.

6.3 Right to object (Article 21 GDPR)

Comment to par. 122–125: We would like to express serious concerns regarding how the right to object under Article 21 GDPR is structured in paragraphs 122–125. The current text introduces broad rejection criteria for data controllers, which are not sufficiently rectified in paragraphs 126 and 127. In fact, paragraph 128 remains the only legal ground for an objection to take actual effect, which becomes highly problematic when coupled with the expansive interpretation of the legitimate interest exception outlined in paragraph 61 and Section 4.3.

When combined with the specific circumstances described in paragraphs 111–120, this framework risks rendering data subjects entirely incapable of exercising their fundamental rights, particularly concerning highly sensitive health data. The current

draft lacks clear, general conditions defining exactly when a data subject's objection must be effectuated by the controller.

A Cautionary Analogy on the Misuse of Legal Exceptions:

The current phrasing of the "legitimate interest" exception creates a significant risk of structural non-compliance. A striking parallel can be drawn with the practical application of Article 5(3) of Regulation (EC) No 261/2004 (Air Passenger Rights), which states that carriers are not obliged to pay compensation if a cancellation is caused by "extraordinary circumstances which could not have been avoided." In everyday practice, this exception has been structurally misused by operating air carriers as a technical workaround to systematically label almost every cancellation as an "extraordinary circumstance," thereby undermining the delivery of passenger rights.

We fear that paragraph 61 of these Guidelines introduces an identical vulnerability. By stating that a controller processing data for scientific research on the basis of legitimate interest can "often attribute significant weight to the processing… in relation to the data subjects' interests," the guidelines open the door for controllers to argue that virtually every data processing activity automatically overrides the data subject's rights.

While patients are theoretically given choices—such as opting out, erasure, or the right to object—any actor seeking to retain data could build a legal and technical workaround based on this broad interpretation, effectively neutralizing those rights. As demonstrated in high-profile data protection cases (e.g., Meta v. Datatilsynet), broad exceptions are highly prone to structural abuse.

Recommendation: It is critical to prevent the systemic misuse of the legitimate interest exemption from the outset through clear guidance, rather than leaving these systemic conflicts to be resolved through protracted litigation. Therefore, we strongly recommend that the EDPB revises this section to provide robust, unambiguous baseline conditions under which a data subject's right to object must be honoured, ensuring that the scientific research exemption does not inadvertently become an absolute barrier to the exercise of fundamental GDPR rights.

7 Attribution of responsibility (controller and processor)

The guidelines correctly follow the classical division of roles under the GDPR, but in a research environment this is often more complex than a clear-cut controller/processor/joint controller model. Research projects typically involve multiple actors (universities, hospitals, infrastructures, funding bodies) who have varying degrees of influence over the processing. In practice, it is not always clear what still counts as "formal participation" and what already constitutes actual determination of purposes or means.

It would be helpful to place more emphasis on a functional approach rather than formal project involvement. At the same time, it would be useful to better explain the relationship to situations where the basic parameters of processing are defined by law. More practical examples from research data ecosystems would significantly improve usability.

7.1 Controller

The definition of controller is correct, but in the research context it is often unclear where the legal or grant limitation ends and where the actual determination of purposes and means begins. In many projects, key processing parameters are pre-set externally, so the space for real decision-making is limited. It is not always clear how these situations affect the determination of the role of controller. It would also help to clarify the role of entities that do not control the processing but set up its framework (e.g. ethics committees or data authorities). The difference between “influencing research” and “determining processing” is difficult to grasp in practice. More concrete examples would increase legal certainty.

7.2 Processor

The Guidelines correctly state that the processor acts on the instructions of the controller, but in research practice it often provides expert services with a certain technical autonomy. However, this does not necessarily mean that it decides the purposes or means of processing. It would be good to confirm more clearly that expert or technical decision-making within the framework of instructions does not change the controller role. The main problems are data platforms, cloud services, or trusted research environments where the technical parameters are to a certain extent standardized. Without further clarification, there is a risk that some technical roles will be incorrectly qualified as controllership. Practical examples would be very helpful here.

7.3 Joint controllers

The Guidelines correctly describe joint controllership, but in a research environment its application may be too broad. Research collaboration should not automatically mean joint determination of personal data processing. In consortia, individual partners often have very different roles and influences on the data. It would help to more clearly separate scientific collaboration from joint decision-making on personal data processing. Otherwise, there is a risk that joint controllership will be used even when it does not correspond to reality. This may unnecessarily increase the compliance burden without a clear benefit for the protection of data subjects.

8 Appropriate safeguards (Article 89(1) of the GDPR)

Chapter 8 is a valuable starting point, but it should more clearly conceptualize appropriate safeguards and data governance across the full research data lifecycle. In data-intensive biomedical, genetic and laboratory research, safeguards should include

not only anonymization, pseudonymization and security measures, but also risk-based access governance, participatory oversight, transparency, auditability, benefit-sharing, management of group harms, governance of clinically relevant findings, and safeguards for downstream uses of data, models and research outputs.

This is particularly important for long-term research infrastructures, biobanks, registries, health data spaces, AI research, genomic and multi-omics datasets, rare disease cohorts and other highly re-identifiable or socially sensitive datasets.

8.1 General

8.2 Risk analysis and data protection impact assessment (DPIA)

Comment to par. 151: Paragraph rightly states that risk analysis or, where required, a DPIA is the starting point for adopting appropriate safeguards. It is also very adequate that the paragraph recognizes that risks in medical research may extend beyond privacy and data protection, including situations where research may affect healthcare provision, for example where a medical condition is identified.

This point should be further developed, especially for biomedical, genetic and laboratory research. In these contexts, research may generate clinically relevant or actionable findings, including genetic findings, incidental findings, secondary findings, or laboratory results that may have implications for the participant's healthcare or for biological relatives. Appropriate safeguards should therefore include governance for the return of results, including criteria for clinical actionability, analytical and clinical validation before any return of individual findings, allocation of professional responsibility, documentation of participants' preferences, respect for the right not to know, and clear interfaces between research pathways and healthcare pathways.

Therefore, we recommend the following formulation: "Where research may generate clinically relevant or actionable findings, the risk analysis or DPIA should address whether and how such findings may be returned to participants or communicated to healthcare professionals. Appropriate safeguards may include governance procedures for the return of results, criteria for clinical actionability, validation before clinical use, allocation of responsibility, documentation of participant preferences, respect for the right not to know, and procedures for cases where findings may also be relevant to biological relatives."

Comment to par. 152 and 153: Paragraphs 152 and 153 appropriately refer to the amount and type of data processed, data linkage, re-identification risks, sharing of research results, vulnerability of data subjects, effects on family members, discrimination and stigmatisation. This is a strong basis, but the Guidelines could more explicitly refer to ELSI risks associated with modern research data governance.

Risk analysis and DPIA should consider not only individual-level risks, but also group harms, community harms, risks to rare disease communities, small cohorts, isolated

populations, minorities and other groups that may be identifiable or affected through research outputs. In addition, data-intensive research increasingly produces derived data, trained models, risk scores, algorithms and other outputs that may create downstream risks even where the original dataset remains access-controlled.

Therefore, we recommend the following wording of the text: Risk analysis and DPIA in scientific research should also consider group-level and community-level harms, including risks of discrimination, stigmatisation, loss of trust, inequitable distribution of risks and benefits, and harms affecting relatives, rare disease communities, small cohorts, isolated populations or minority groups. In data-intensive research, the assessment should also cover derived data, trained models, algorithms, risk scores and downstream reuse of research outputs.

8.3 Anonymisation and pseudonymisation of personal data

Comment to par. 157-160: Paragraphs 157–160 correctly explain that pseudonymisation should be used where anonymised data cannot fulfil the research purposes, for example where individual-level follow-up over time is needed. In biomedical, genetic and laboratory research, this clarification is particularly important.

The Guidelines should make clear that pseudonymisation is not merely a substitute solution where anonymisation is impossible. In many biomedical contexts, pseudonymisation is the most appropriate safeguard because it enables longitudinal follow-up, validation of research results, reproducibility, quality control, traceability of biological samples, linkage with clinical records or registries, and, where appropriate, controlled re-contact or return of clinically relevant findings. Treating anonymisation as the default without sufficient recognition of these needs may create tension with scientific validity, clinical safety and responsible research governance.

8.4 Safeguards when processing genetic or biometric data

Comment to par. 165-168: Section 8.4 rightly recognises the risks of genetic and biometric data, including their relevance for relatives, health predispositions, inherited diseases and lifelong identifiability. This section should be expanded to explicitly address genomic and omics research, including whole-genome sequencing, whole-exome sequencing, multi-omics datasets, rare disease cohorts, small cohorts and highly granular phenotype-genotype datasets.

These data may remain highly re-identifiable even after removal of direct identifiers. They may also have implications for biological relatives, family structures, reproductive choices, insurability, employability, stigma and discrimination. The Guidelines should therefore recommend more specific safeguards for high-risk genomic and omics datasets, including secure processing environments, federated analysis, restrictive access governance, independent review of access requests, limitations on local

download, strong publication controls, prohibition of unauthorised re-identification, and specific rules on return of results.

Comment to par. 169: Paragraph 169 recognises the need for information campaigns in genetic research involving isolated population groups. Similar concerns may arise not only in isolated populations, but also in rare disease communities, founder populations, small national or regional cohorts, minority groups, indigenous communities and other groups that may be identifiable or affected by genetic or genomic research.

The Guidelines should also clarify that transparency should not be limited to one-way information campaigns. In higher-risk contexts, participatory governance may itself function as an appropriate safeguard. This may include patient and public involvement, community advisory boards, representation of affected groups in governance bodies, benefit-sharing arrangements, feedback of aggregate results, and mechanisms to address community concerns.

8.5 Determination of other types of appropriate safeguards

Comment to par. 170: Paragraph 170 includes governance structures, review boards, oversight committees and representation by data subjects as examples of safeguards. However, the Guidelines should provide more concrete guidance on access governance for research data infrastructures.

For biobanks, registries, health data repositories, federated platforms and long-term research infrastructures, safeguards should include transparent data access criteria, data access committees, review of scientific and ethical merit, assessment of commercial uses, restrictions on onward sharing, audit logs, publication rules, sanctions for misuse, and periodic review of approved access. These measures are central to responsible ELSI-informed data governance and should not be treated as optional or peripheral.