



## European Network of Expertise (NoE) on Omics in Cancer

(Joint Action JANE-2 – GA 101183265 – WP9)

**Contribution to the EDPB Guideline entitled**

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

**Key issues and recommendations**

**25<sup>th</sup> June 2026**

### **What is JANE-2?**

JANE-2 (Joint Action Networks of Expertise in Cancer, <https://jane-2.eu/>) is an ambitious initiative, stemming from Europe’s Beating Cancer Plan, with the aim to implement seven new European Networks of Expertise in different cancer conditions, addressing: 1) complex and poor prognosis cancers; 2) palliative care; 3) survivorship; 4) personalized primary and secondary prevention; 5) omics technologies, 6) hi-tech medical resources; and 7) adolescents and young adults with cancer. It represents 121 partners across 29 European countries.

### **Why a Network of Expertise on Omics Technologies?**

Omics technologies are revolutionizing the management of patients with cancers, haematological malignancies, as well as rare diseases. These innovative tools go beyond supporting precision diagnosis and treatment,. They also improve cancer prediction, early detection, and disease monitoring. Genomics and transcriptomics are currently the most widely used in clinical practice. However, emerging fields such as methylomics, proteomics, and metabolomics, as well as multi-omics integration, are rapidly advancing and poised to play a significant role in the future.

The ambition of the Network of Expertise on Omics (NoE) (<https://jane-2.eu/omics-technologies/>) is to support the integration of innovative omics technologies into the standard of care at the different steps of cancer management in a sound and sustainable manner, achieving equitable access to these services for all EU citizens.

### **Addressing unmet urgencies and barriers relevant to omics-driven diagnostics within the GDPR framework**

Within the broader context of GDPR implementation for scientific research, it is also important to recognise that many omics-driven diagnostic activities rely on in-house analytical pipelines, laboratory-developed methods, and institution-specific validation processes that evolve continuously as scientific knowledge advances. These diagnostic workflows are integral to health research and often represent the first point at which new biomarkers, signatures, or multi-omics patterns are identified,



validated, and translated into clinical insight. When data-protection requirements are interpreted in ways that are overly rigid, duplicative, or disconnected from established quality frameworks already applied in hospital and research laboratories, they can inadvertently slow down diagnostic innovation, limit the development of evidence-based tests, and delay access to timely solutions for patients, particularly in rare, complex, or fast-evolving disease areas.

Health institutions and academic laboratories already operate under stringent scientific and quality standards, including ISO-based accreditation systems, ethical oversight, and robust governance structures for omics data. These frameworks ensure analytical validity, traceability, and patient safety. GDPR implementation should therefore complement, not duplicate or constrain, these existing safeguards. A proportionate, research-aware interpretation of GDPR obligations is essential to ensure that laboratories can continue to develop and refine omics-based diagnostic approaches, share expertise within clinical and research networks, and contribute to the rapid generation of high-quality evidence.

This is particularly important in cross-border research, where alignment of expectations, mutual recognition of safeguards, and coherent interpretation of GDPR provisions are indispensable. Fragmented or inconsistent requirements across Member States create unnecessary barriers to collaboration, hinder the exchange of expertise, and complicate the operation of multi-country research infrastructures. A context-sensitive, risk-based, and proportionate approach, one that recognises the scientific justification for processing and the safeguards already in place, is essential to support meaningful collaboration while maintaining strong protection for individuals. Above all, GDPR guidance must be coherent, predictable, and suitable for the realities of modern omics research so that institutions developing omics technologies across Europe can work together effectively and responsibly.

## General comments

The EDPB Guidelines 1/2026 represent an important step toward clarifying the application of the GDPR to scientific research. However, when examined against the operational realities of omics research, clinical studies, long-term cohort infrastructures, and emerging EU data-governance frameworks, several structural gaps remain. These gaps are not marginal: they directly affect the feasibility, legal certainty, and cross-border continuity of genomics and multi-omics research across the Union. They also risk undermining coherence between the GDPR and the rapidly evolving EU health-data ecosystem, including the EHDS, DGA, CTR, IVDR/MDR, and the AI Act, all of which are central to the responsible generation, linkage, and reuse of omics data. Across Member States, researchers, ethics committees, and data-protection authorities face divergent interpretations of core GDPR concepts: purpose, legal basis, pseudonymisation, compatibility, controllership, and data-subject rights. These divergences create fragmentation that directly impedes scientific progress, delays approvals, and generates inconsistent expectations for controllers. The Guidelines, while conceptually robust, do not yet provide the operational clarity needed to ensure harmonised implementation in complex research settings.

As an overall comment on the Guidelines, we observe that current EU governance frameworks for (gen)omics data do not yet fully reflect the realities of contemporary science or the needs of real-world research. Genetic and omics information is still largely regulated as if it were a purely individual attribute, something that belongs to one person, sits in one file, and carries a single, isolated set of risks. In practice, this is no longer the case. Today, (gen)omics data is simultaneously personal, relational, and relevant at broader societal levels. A single dataset can inform clinical decisions for an individual, provide insights for families and communities, and contribute to the strategic capabilities of health systems. Yet our legal frameworks remain anchored in an individual-centric logic that does not fully accommodate these overlapping dimensions. This misalignment has consequences :

- it slows down research that could save lives,
- it complicates cross-border collaboration,
- it fragments or delays the set up and governance of omics infrastructures,
- and it leaves policymakers, researchers and innovators navigating contradictions instead of clarity.

What we need now is a governance framework that **recognises the layered nature of (gen)omics data** and supports decision-making across all dimensions: individual rights, collective implications, and strategic societal interests. A framework that is proportionate, coherent, and capable of enabling innovation to reach people *when it matters*, not years later.

**Several aspects that require further attention emerge clearly from our analysis:**

1. **Persistent fragmentation across Member States.** Divergent national laws, different interpretations of Data Protection Authorities, and ethics committee requirements continue to

produce inconsistent compliance expectations for the scientific research communities. This fragmentation is particularly acute in areas such as secondary use, pseudonymization requirements, legal bases, and the exercise of data subject rights in multi-actor research chains.

2. **Lack of operationalisation of the GDPR applied to health and scientific research.** The Guidelines should move beyond abstract principles and provide practical, operational guidance aligned with existing scientific, ethical, and regulatory standards (e.g., ICH GCP, CTR, IVDR/MDR, EHDS). GDPR bodies should not redefine research concepts but ensure GDPR is applied coherently within established research frameworks. A coherent European data-governance framework should recognise the specificity of health research, rely on already established sectoral standards, and encourage stakeholders to develop harmonised principles for data protection in research in the same spirit as internationally recognised frameworks such as International Conference Harmonisation – Good Clinical Practice (ICH-GCP, <https://www.ich.org/>). We recommend creating Good Data Protection Practices or Code of Conduct, applied to Health Research.
3. **Lack of clarity on the interplay between the GDPR and other EU legal instruments.** The Guidelines do not sufficiently articulate how GDPR provisions interact with the CTR, IVDR/MDR, EHDS Regulation, DGA, or AI Act. As a result, controllers face overlapping, sometimes contradictory obligations, especially in health research ecosystems where multiple legal frameworks apply simultaneously.
4. **Provide a harmonised, research-specific interpretation of “purpose”.** Scientific research within a defined field, infrastructure, disease area, should be treated as one overarching purpose, rather than fragmented into separate purposes for each project, dataset, infrastructure or analytical step. A harmonised, research-specific interpretation of purpose is essential to reduce fragmentation, avoid institutional inertia, and ensure proportionate governance across Member States. Moreover, clarifying purpose must be accompanied by a clearer articulation of roles and responsibilities, particularly around controllership. In practice, controllership is purpose-specific, not permanent, and the transitions in responsibility, especially under the EHDS Regulation, between data holder, HDAB, and data user during a secondary-use request, remain a source of operational uncertainty. While distinct from the definition of purpose, these role allocations are tightly interlinked: without clarity on both, institutions default to risk-averse interpretations that hinder legitimate research and complicate the governance of health data in general and omics data in particular. A research-appropriate framing of purpose, coupled with explicit guidance on role boundaries and the permissibility of maintaining context-dependent safeguards, controlled access-oriented would support continuity, reuse, and long-term stewardship of health data while maintaining strong data-subject protection.
5. **The Guidelines should explicitly acknowledge that GDPR consent-based models are not feasible** as the primary legal basis for long-term reuse of data, omics research infrastructures or research on linked data from several sources. The GDPR already provides appropriate alternatives, such as Article 6(1)(e) together with Articles 9(2)(i) or 9(2)(j), supported by Article 89 safeguards; and the Guidelines should clarify that broad or dynamic consent is an optional enhancement where feasible, not a baseline expectation. Explicit recognition of these structural limits would reduce

fragmentation, promote consistent interpretation across Member States, and safeguard the continuity of research. At the same time, the Guidelines should acknowledge that new studies launched today will fall under the future secondary use framework of the European Health Data Space Regulation, which may introduce obligations that appear to contradict current expectations around “future processing” in consent forms. Clear, forward-looking guidance is therefore needed on what information must be included in consent-based studies so that participants are not misled, and researchers are not forced to revise consent materials repeatedly as EU law evolves.

6. **Clarify the distinction between GDPR consent and ethical consent without imposing dual tick-box systems that confuse patients.** While the distinction between GDPR consent and ethical/Helsinki consent may be clear to professionals, it is not meaningful or intuitive for patients, especially for vulnerable populations. Participation in research by design requires the processing of personal data, and introducing two separate tick boxes (one for study participation and one for GDPR consent) creates a high risk of inconsistent choices. A patient may agree to participate in the study but inadvertently decline data processing, making full participation impossible. This approach is unreliable, confusing, and contrary to a patient-centred, accessible consent process. To avoid this contradiction, the Guidelines should explicitly allow a single, integrated consent statement that reflects the operational reality of research. A clear and patient-friendly formulation could be: *“If you agree to participate in this study, you also accept that your data will be processed for the purposes of this research.”* This preserves transparency, respects autonomy, and avoids unnecessary administrative complexity, while ensuring that participants are not misled by artificial distinctions that do not reflect how research is conducted in practice.
7. **Recognise the interdependence of research and infrastructure.** Acknowledge that research and the infrastructures enabling it, registries, biobanks, secure platforms, and federated environments, are not separate purposes but operationally inseparable components of lawful scientific research. These infrastructures are not merely technical service providers: they are scientifically governed, security and technical compliant, GDPR-compliant expert centres that ensure data quality, continuity, linkage, and controlled re-identification where scientifically necessary. They should therefore be treated, financed, and evaluated as integral parts of the research ecosystem, not as external utilities. Their long-term operation, including the maintenance of structural linkages and enriched cohorts, falls squarely within the overarching research purpose and should not be subject to repeated deletion or reconstruction cycles tied to individual project-level purposes. Recognising infrastructures as embedded research actors, rather than separate entities, supports data quality, scientific validity, regulatory compliance, and the sustainability of high-quality health research.
8. **Clarify the cumulative application of Articles 6 and 9 GDPR.** Explain how Articles 6 and 9 operate together in light of the EHDS, and ensure that combinations of legal bases do not circumvent EHDS safeguards for AI training, testing, and evaluation.
9. **Unresolved complexities in research roles and responsibilities.** The Guidelines do not address the practical reality that investigators often “wear different hats” (care vs research), that sponsor–investigator relationships vary across Member States, or that public health institutes may act

simultaneously as data holders, data users, and statutory authorities. These ambiguities hinder consistent role allocation and accountability.

10. **Insufficient operational guidance on data subject rights in research.** While the Guidelines mention restrictions to erasure and objection, they do not explain how rights should be exercised when sponsors hold only pseudonymized data, when multiple controllers are involved, or when re-identification is neither permitted nor possible. This gap leads to unrealistic expectations and inconsistent practices across Member States.
11. **Lack of a risk-graded approach to omics data.** The Guidelines treat all genetic data as a single category, despite the wide risk spectrum between WGS/WES, tumour only somatic profiles, and single biomarkers. This “one size fits all” approach leads to disproportionate safeguards, unnecessary barriers to research, and confusion for ethics committees and DPAs.
12. **Absence of research-tailored guidance on international data transfers.** Provide research-tailored guidance on international data transfers. Current transfer tools, SCCs, TIAs, and ad hoc clauses, are not fit for multi-actor, long-term scientific health research. Publicly funded and academic organisations often cannot conduct individual Transfer Impact Assessments (TIAs) or sign the existing Standard Contractual Clauses (SCCs) in their commercial form, yet rely on cloud-based analytics, federated models, and multi-year processing of pseudonymised data. Existing transfer instruments are also unstable, frequently challenged or revised, creating uncertainty for infrastructures that must guarantee continuity and reproducibility. We therefore strongly recommend that the EDPB establish research-appropriate transfer routes that allow lawful sharing of health data, especially pseudonymised data, when SCCs or adequacy decisions are unavailable, legally impossible, or disproportionate. This should include solutions for academic and public bodies unable to sign SCCs, secure cloud processing, and federated data-visiting models. Such routes are essential to ensure that high-quality, long-term research can operate without recurring legal and administrative dead-ends.
13. **Given the consistent interpretation in the health sector, and now explicitly in the European Health Data Space Regulation,** that the processing of personal electronic health data requires a cumulative legal basis under both Article 6 and Article 9 GDPR, we would kindly ask the EDPB to clarify this interplay in the forthcoming Guidelines on scientific research. In particular, it would be helpful for the EDPB to explain how Articles 6 and 9 should be applied in light of the EHDS provisions (e.g., Articles 53, 67(4), 68(1)(c)) and whether combinations such as Article 6(1)(f) with Article 9(2)(k) could inadvertently circumvent the safeguards, governance mechanisms and access conditions established under the EHDS for AI training, testing and evaluating algorithms in scientific research questions. Such clarification would support consistent interpretation across Member States and ensure alignment between GDPR guidance and the new EHDS framework.
14. **It would be highly valuable for the EDPB to include a dedicated section recommending mandatory training modules** for stakeholders involved in interpreting the GDPR in the context of health research. Specifically, training for ethics committees on GDPR principles and for Data Protection Authorities on clinical research and sector-specific frameworks.



Taken together, these issues demonstrate that while the Guidelines provide a strong conceptual foundation, they require further development to ensure harmonised, feasible, and legally certain implementation across the EU health research landscape. Our recommendations aim to support the EDPB in strengthening the Guidelines so that they fully reflect the realities of modern scientific research, enable responsible data reuse, and uphold the fundamental rights of individuals while safeguarding Europe's capacity for high-quality, cross-border research.

We also note that recent documents issued by EU institutions sometimes offer recommendations that are not fully aligned, particularly regarding the balance between competitiveness, the reuse of health data for research, and the paramount importance of data protection. Research communities are committed to compliance, but they require consistent and sustainable guidance to navigate data-governance obligations effectively. The current fragmentation risks creating uncertainty for controllers, processors, and research infrastructures, especially in Member States preparing for EHDS implementation. Greater alignment between the GDPR and the sector-specific frameworks is essential to ensure legal certainty and operational feasibility.

We appreciate the EDPB's efforts to clarify the application of the GDPR to scientific research and support the overarching aim of enabling responsible, high-quality research while protecting fundamental rights. To maximise the Guidelines' impact, we encourage:

- further alignment with parallel EU initiatives,
- clearer operational guidance for research and health institutions, and
- explicit recognition of the interdependence between primary and secondary use in health data ecosystems.

In preparing our submission, we have reviewed each section of the Guidelines and highlighted both elements that are clear and directly usable, and areas where further development or operational guidance is needed.

## Section-specific gaps

Scientific research in the health field, and especially omics research, operates in environments fundamentally different from other sectors' data processing. Omics ecosystems depend on multi-actor research chains that link diverse data types, long-term infrastructures, and high-risk, high-value data categories such as genetic and other omics data.

These research domains also rely on hospital-collected data, registries and biobanks, where data are gathered continuously, often without direct contact with individuals, and are governed by different legal frameworks<sup>1</sup>.

Secondary use of health data is essential for scientific research, progress and for the timely development of solutions that benefit patients, particularly in hospitals where innovation begins and clinical practice meets research. Health research, including genomics and multi-omics, depends on the ability of healthcare professionals to work with real-time data and patient samples so they can explore hypotheses, validate observations, and drive meaningful innovation. These activities are fundamental to improving diagnosis, treatment and prevention, and they rely on data environments that support responsible reuse while maintaining strong protections for individuals. **It is therefore crucial that legal and ethical frameworks enable such research to take place, offering coherence, adaptability and flexibility, and above all, doing so in a timely and sustainable manner that supports the realities of modern health research.**

Health data in general and omics in particular, is biologically shared, analytically interdependent, and foundational for next-generation diagnostics and personalised medicine. Treating it as if it were simply another category of individual-level data processed by a single controller ignores how omics infrastructures actually function.

This is why a sector-specific, omics-aware governance approach is essential: to ensure that GDPR principles are applied in a way that is proportionate, harmonised, and scientifically feasible, while still protecting individuals' rights.

Without tailored guidance, Member States, DPAs, and ethics committees will continue to interpret key concepts inconsistently, creating fragmentation that directly impedes cross-border omics research and undermines Europe's ability to generate high-quality evidence for public health and clinical innovation.

### 1. Strengthening the six key-indicative factors through omics research examples

#### Issue statement

A central weakness in the current Guidelines is that the six key indicative factors for determining whether processing qualifies as scientific research remain too abstract to guide real-world decision-making in omics-driven health research, innovation-driven omics research, and advancements in personalised medicine overall. While conceptually sound, the factors are presented at a level of generality that does not reflect the operational realities of research ecosystems that function at individual and population scales, and evolve over decades.

---

1

[20250401 study on the secondary use of personal data in the context of scientific research 23102020\\_en.pdf](#)

We do not believe that the definition of scientific research should be fragmented across multiple EU regulations. Instead, it should be governed by clear, stable principles developed by multidisciplinary teams with expertise in health research. Such principles already exist in internationally recognised frameworks, such as the Helsinki Declaration, the Nuremberg Code, the Taipei Declaration, the Oviedo Convention, and guidance from WHO and OECD, which collectively articulate the ethical and scientific foundations of research.

However, if the EDPB intends to further elaborate criteria or provide examples, this work should build on these established principles rather than create parallel or conflicting definitions. In that context, the Guidelines should include omics-enabled research domains, such as but not limited to omic data analysis for personalised medicine, national disease registries, and omics infrastructures, which depend on continuous data collection, multi-actor governance chains, and secondary use as an essential scientific function, not an optional add-on. These infrastructures often operate without direct contact with individuals, and recontact is frequently impossible, disproportionate, or in some instances ethically inappropriate. Yet the Guidelines implicitly assume project-based, short-term, individually consented research models that do not map onto the realities of modern omics ecosystems.

The absence of concrete, domain-specific examples creates significant interpretative gaps across Member States. Without illustrations that reflect the operational realities of omics-intensive applications, the six factors are left to apply inconsistently. This inconsistency is already visible: divergent readings of what qualifies as “scientific research” hinder secondary use, delay approvals, and leave controllers uncertain about how to demonstrate compliance. In some jurisdictions, the factors are applied so narrowly that long-term or population-level omics research risks being excluded altogether; in others, they are interpreted so broadly that they lose practical value. Both tendencies undermine harmonisation and weaken Europe’s capacity to generate high-quality evidence for personalised medicine, public health, and clinical innovation.

### Need for clearer articulation in the Guidelines

To ensure that the six factors function as usable tools rather than abstract principles, the Guidelines should include operationalised examples that translate each factor into contexts where scientific research depends on continuity and scale. Examples drawn from genomic research at the population level, national omics research projects, and longitudinal multi-omics cohorts would give controllers concrete guidance, support consistent assessments by Health Data Access Bodies under the EHDS, and ensure that population-level and long-term omics research is clearly recognised within the GDPR framework.

The examples below are merely illustrative. In reality, these challenges are shared across all research organisations.

| Key indicative factor                 | Example   | Description   |
|---------------------------------------|---|---|
| 1. Methodical and systematic approach | <b>A. National cancer registry + whole-genome linkage study</b> | A national public health institute links cancer registry data with whole-genome sequencing datasets to evaluate 10-year survival trends, following a predefined analysis plan, registry linkage protocol, and quality-control procedures. |
|                                       | <b>B. Longitudinal multi-omics birth cohort</b>                 | A cohort established in 2010 continues to collect genomic, epigenomic, metabolomic, and environmental exposure data. Despite evolving hypotheses, the research follows a structured protocol,   |

|   |  |  |
|---|--|--|
|   |  | governance framework, and documented data-collection methodology.  |
|   | <b>C. Genomic infectious-disease surveillance</b>                          | Weekly sequencing data from sentinel laboratories are analysed using a harmonised epidemiological protocol to detect emerging variants and unusual increases in pathogen circulation.                            |
| <b>2. Adherence to ethical standards</b>          | <b>A. Secondary use of stored omics biosamples</b>                         | A biobank established for cancer research allows secondary use of stored serum and tumour samples for multi-omics profiling under ethics committee oversight and national biobank legislation.                   |
|   | <b>B. AMR genomic surveillance with safeguards</b>                         | A public health authority conducts antimicrobial-resistance surveillance using pseudonymised genomic data, overseen by an independent ethics board ensuring proportionality and risk minimisation.               |
|   | <b>C. Historical omics data where recontact is impossible</b>              | A study uses archived sequencing data from the 1990s to analyse long-term cancer risk. Ethical review confirms recontact is impossible and that safeguards (secure environments, pseudonymisation) are in place. |
| <b>3. Verifiability and transparency</b>          | <b>A. Publication of cohort and omics methodology</b>                      | A longitudinal cohort publishes its recruitment strategy, omics pipelines, data dictionary, and statistical methods, enabling reproducibility.   |
|   | <b>B. Open genomic surveillance reports</b>                                | A national institute publishes weekly genomic surveillance bulletins with sequencing methods, case definitions, and data sources.  |
|   | <b>C. Peer-reviewed validation of omics-based algorithms</b>               | A research team developing an AI-based early-warning model for sepsis publishes the model architecture, validation metrics, and limitations.   |
| <b>4. Autonomy and independence</b>               | <b>A. Independent public health omics analysis</b>                         | A national institute analyses vaccine effectiveness using genomic data, independent of the manufacturer that supplied the product.   |
|   | <b>B. Academic–industry collaboration with independent scientific lead</b> | A university leads a pharmaco-omics study using real-world data; although industry provides funding, the academic team independently defines the research questions and publication plan.                        |
|   | <b>C. Independent governance of multi-omics cohorts</b>                    | A cohort’s scientific committee, composed of independent researchers, approves data-access requests and ensures analyses are not influenced by funders.  |
| <b>5. Contribution to generalisable knowledge</b> | <b>A. Long-term multi-omics trends in chronic disease</b>                  | A 30-year cohort study identifies genomic and environmental risk factors for diabetes and cardiovascular disease, producing generalisable findings applicable to EU populations.                                 |
|   | <b>B. Genomic surveillance informing public health policy</b>              | Analysis of national sequencing data identifies regional disparities in variant circulation, informing targeted interventions.   |
|   | <b>C. Environmental exposure + omics research</b>                          | A study linking air-quality data with transcriptomic and clinical data generates evidence relevant for environmental health policies.  |
| <b>6. Dissemination of results</b>                | <b>A. Annual omics-integrated epidemiological reports</b>                  | A public health authority publishes annual reports on cancer incidence, AMR genomics, or chronic disease trends.   |
|   | <b>B. Cohort omics data sharing for secondary research</b>                 | A cohort makes pseudonymised multi-omics datasets available to external researchers under controlled access.   |

|  |   |   |
|--|---|---|
|  | <p><b>C. Public dashboards and open summaries</b></p> | <p>A national institute provides aggregated, anonymised omics indicators (e.g., variant trends, environmental exposures) through public dashboards.</p> |
|--|---|---|

## 2. Persistent ambiguity around the definition of “purpose” in research creates structural barriers to lawful secondary use

### Issue statement

Does the GDPR term “purpose” have a different granularity in the scope of scientific research as compared to other sectors? In the field of human resources, the purpose would usually be formulated on a relatively high level, such as “processing of personal data for execution of the employment contract”. This purpose will be achieved through different processing activities (payment of salaries, compliance with social laws etc...). Similarly, in marketing, the purpose would be defined as “using personal data for the marketing purpose”. In the field of research, however, some countries and organisations are considering the infrastructure hosting the health data, and each research project as different purpose, including when research projects are very similar. This approach is a clear overshoot in our view. A clarification on the acceptable granularity of the purpose would be welcome. For example, can the purpose be formulated as follows: “building databases, maintaining and (re)using such health data for cancer research and related diseases”? We believe that such an interpretation will be consistent with the EDPB Opinion 3/2019<sup>2</sup>, which sees clinical trials as processing activity, and not as purpose (the purpose would be series of clinical trials performed for the similar purpose indeed).

A major unresolved problem in the current regulatory landscape is the absence of a clear, operational definition of “purpose” as it applies to scientific research. The Guidelines acknowledge that research may evolve, yet they do not provide concrete criteria for determining when different research activities fall within the same overarching purpose and when they constitute distinct purposes. This gap has led to divergent national interpretations, with some authorities and institutions treating every new research project, even those within the same disease area, methodology, or long-term programme, as a separate purpose under the GDPR. This approach is not only conceptually misaligned with how “purpose” is understood in other domains (such as employment or marketing), but also creates unnecessary operational and legal burdens for researchers.

The lack of clarity has significant downstream effects. When each project is treated as a new purpose, controllers are pushed toward repeated re-consenting exercises, even when the underlying scientific aim remains consistent and the data continue to be processed within the same broad research framework. This dynamic fuels the very fragmentation and uncertainty that the GDPR sought to avoid. It also undermines the feasibility of long-term research infrastructures, where secondary use is essential and where re-contacting participants is often impossible, disproportionate, or ethically inappropriate. Without a coherent definition of purpose, controllers face inconsistent expectations across Member States, Ethics Committees impose requirements that exceed their remit, and secondary use becomes entangled in procedural obstacles that do not enhance data protection but do impede scientific progress.

**What is missing is a principled, harmonised articulation of how “purpose” should be interpreted in the research context, one that recognises the continuity of scientific inquiry, the legitimacy of broad research aims, and the**

<sup>2</sup> [https://www.edpb.europa.eu/sites/default/files/files/file1/edpb\\_opinionctrq\\_a\\_final\\_en.pdf](https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_opinionctrq_a_final_en.pdf)



**safeguards already embedded in Article 89.** Until this is clarified, purpose fragmentation will continue to generate avoidable re-consent demands, legal uncertainty, and operational delays, ultimately weakening Europe’s capacity to conduct high-quality, cross-border research.

The EDPB guideline does not clarify how “purpose” should be interpreted for **long-term, evolving, multi-purpose health research**, especially when:

- rare-disease datasets cannot be anonymised without destroying utility (“true anonymisation practically impossible”).
- research requires **repeated reinterpretation** of genetic/omic data over time.
- secondary use is essential for scientific validity (“significant uncertainty remains regarding the rules that apply to the reuse...”).

**The practice of classifying every project as a new purpose is an excessive and unjustified expansion of the GDPR’s purpose limitation principle.**

#### Need for clearer articulation in the Guidelines

Would be very welcome that the Guideline clarifies:

- Whether building health databases (through designed registries or simply collected as part of a defined clinical research), maintaining and making available for (re)use as a research project constitutes a separate purpose. We strongly believe that this should not be the case and we support this with argumentation at below point 11.
- Whether secondary use within the same disease area is one purpose.
- How to avoid unnecessary re-consenting due to purpose fragmentation.

#### Concrete proposals

To avoid fragmentation and unnecessary re-consenting, the EDPB should clarify that **“scientific research” constitutes a broad, overarching purpose under the GDPR**, within which multiple research projects, analyses, and follow-up studies may take place. This interpretation is consistent with Recital 159 GDPR, the presumption of compatibility, and the operational realities of omics research.

Proposed way forward:

- Avoid granular, activity-by-activity (purpose granularity) qualification of controller/processor roles.
- Encourage a broad, research-purpose-based approach, aligned with how research is actually conducted.
- Recognise population-level research as a continuing purpose rather than a series of discrete purposes.
- Allow controllers to define a purpose cluster (e.g., surveillance, epidemiology, variant reinterpretation, method validation, broad research question).
- Clarify that reinterpretation of genomic data over time is compatible with the original purpose when safeguards are in place.
- Provide examples for rare diseases where purpose cannot be narrowed without undermining scientific value.

We propose wording for the Guidelines:

*“For the purposes of Article 5(1)(b), scientific research should be understood as a broad, overarching purpose. Research activities conducted within the same research infrastructure, disease area, or public-health mandate should not be considered new purposes requiring re-consent, provided that appropriate safeguards under Article 89(1) are in place, the research question is ethically sound and the processing remains compatible with the expectations established at the time of data collection.”*

*“Re-consent should not be required where recontacting individuals is impossible or disproportionate, where the research remains within the same overarching scientific purpose, and where ethical oversight and robust safeguards are applied.”*

Why this solution is necessary:

- Prevents fragmentation across Member States.
- Avoids unnecessary re-consenting that is impossible in long-term or population-level research.
- Aligns GDPR with EHDS, CTR, and public-health mandates.
- Supports continuity of registries, cohorts, and biobanks.
- Reflects scientific reality: research evolves, but the purpose often remains the same.

#### Concrete examples from health research

| Topic  | Proposed clarification  | Rationale  | Example   |
|--|---|--|---|
| <b>1. Not every research project is a separate purpose</b>                             | <p>A new research project does <b>not</b> constitute a new purpose when it:</p> <ul style="list-style-type: none"> <li>• falls within the same overarching scientific aim,</li> <li>• uses data from the same research infrastructure (cohort, registry, biobank),</li> <li>• is conducted under the same governance and safeguards,</li> <li>• remains compatible with expectations at initial data collection.</li> </ul> | Prevents artificial fragmentation of “purposes” in long-term research ecosystems; aligns with scientific practice where multiple analyses serve one overarching aim. | A cancer cohort established to study long-term health outcomes performs multiple analyses over 20 years (e.g., cancer incidence, environmental exposures). These are <b>sub-purposes</b> , not new purposes.  |
| <b>2. Secondary use within the same disease area or research domain is one purpose</b> | <p>Secondary use of data within the same disease area, research domain, or public health mandate should be treated as <b>one purpose</b>, provided safeguards under Art. 89(1) are in place.</p>  | Ensures continuity of disease-specific research; avoids unnecessary administrative burden; supports public health monitoring.  | A national cancer registry uses the same dataset to study survival, treatment effectiveness, late effects, and regional disparities. These fall under the overarching purpose of <b>cancer epidemiology</b> . |

|   |   |   |   |
|---|---|---|---|
| <p><b>3. Mechanism to avoid unnecessary re-consenting</b></p> | <p>Re-consent should <b>not</b> be required when:</p> <ul style="list-style-type: none"> <li>• the new activity is compatible with the original purpose,</li> <li>• recontacting individuals is impossible or disproportionate,</li> <li>• research is under ethical oversight,</li> <li>• safeguards (14harmonized, secure environments) are applied.</li> </ul> | <p>Aligns with Art. 5(1)(b), Art. 89(1), Recital 33 (broad consent), Recital 50 (compatibility), and Helsinki Art. 32. Supports the feasibility of long-term and population-level research.</p> | <p>A biobank collected samples in 1998 for cancer research. A 2026 genomics study aims to identify hereditary risk factors. Re-contact is impossible. Under Art. 89(1) and Recital 33, <b>re-consent is not required.</b></p> |
|---|---|---|---|

### 3. Overestimate the feasibility of consent-based models in long-term, population-level research infrastructures

#### Issue statement

One positive contribution of the Opinion No 3/2019<sup>3</sup> by the EDPB consists in stressing out the fact that the obligation for informed consent pursuant to clinical research should be distinguished from consent as one of the possible grounds for the processing of personal data under the GDPR. As highlighted in other reports on Opinion 3/2019, presenting subjects with multiple consents that comply with different laws could be confusing for clinical trials participants<sup>4</sup>, hence special attention must be paid to the communication between sponsors/investigators and patients in this regard. Moreover, the EDPB and the EU Commission show a clear preference towards the use of other legal bases for the processing of personal data in the context of clinical trial (namely, public interest or the legitimate interest of the controller) rather than relying on consent. And the limits of consent as a legal basis for processing data in clinical research were clearly highlighted (not granular enough, not feasible for research, patient population is vulnerable, consent requires active yes/no tick-boxes, etc..).

However, several crucial topics that are related to the matters discussed in the documents, have been omitted. Among those falls the importance of the One-stop-shop mechanism (hereinafter “OSS”) for pan-European clinical research. More specifically, the influence that the choice of the legal basis for the processing of personal data has over whether OSS will apply, or not. In our view, it would be a welcomed development if the EDPB choses to issue further guidance on the issues pertaining to the OSS and the choice of the legal basis for international research activities. The GDPR is lauded for introducing the OSS. The mechanism’s rationale is that controllers who conduct cross-border processing of personal data<sup>5</sup> benefit from dealing with a single point of contact (lead supervisory authority<sup>6</sup>) within the EU.

Based on further guidelines for identifying a controller or processor’s leading supervisory authority, it is clear, a clinical research open in several Member states, would fall within the scope of these provisions. This reflects the

<sup>3</sup> [https://www.edpb.europa.eu/sites/default/files/files/file1/edpb\\_opinionctrq\\_a\\_final\\_en.pdf](https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_opinionctrq_a_final_en.pdf)

<sup>4</sup> [EDPB issues new opinion on interplay between Clinical Trials Regulation and the GDPR | Data Protection Report](#)

<sup>5</sup> Article 4(23) GDPR defines cross-border processing of personal data as either “processing of personal data which takes place in the context of the activities of establishments in more than one Member state of a controller or processor in the Union, where the controller or processor is established in more than one Member state”, or “processing of personal data which takes place in the context of the activities of a single establishment of a controller or processor in the Union but which substantially affects or is likely to affect data subjects in more than one Member state”.

<sup>6</sup> Article 56 GDPR defines it as the Supervisory authority of the main establishment or of the single establishment of the controller or processor.

unification and simplification which are proclaimed to be at the heart of the data protection reform. The involvement of several data protection authorities (DPAs) may generally lead to confusion for controllers, conflicts of competence, and uncertainty for data subjects. When analysing the possible legal bases for data processing in the context of a clinical research, of major importance is Recital 128 GDPR, pursuant to which the rules on OSS do not apply where the processing is carried out in the public interest. Further to the Recital 128 GDPR, art 55.2 specify that the notion of the lead supervisory authority (specified in the art.56) does not apply where the processing is carried out in the public interest or when processing is necessary for compliance with a legal obligation. So, choosing the public health or compliance with legal obligation as the legal basis in international led clinical research comes with challenges. First, each member state can further specify conditions which allows to rely on this legal basis, which may not be possible to meet in all concerned countries. Secondly, such legal basis is not compatible with the OSS mechanism.

This means that the data controller would not be able to benefit from a lead supervisory authority in the following cases:

- i) Processing of personal data for reliability and safety purposes, where the correct legal basis has been identified unanimously by EDPB and the EU Commission as Article 6(1)(c) in conjunction with Article 9(2)(i) of the GDPR;
- ii) Processing operations related to research activities, if one of the following legal bases is chosen by the controller: Article 6(1)(e) in conjunction with either Article 9(2)(i) or (j).

**It follows from the foregoing that the legal basis that is left to be the most preferable, is Article 6(1)(f) (legitimate interest) in conjunction with Article 9(2)(j) (scientific research).** By choosing it, the data controller would have access to the OSS mechanism concerning the processing of personal data when it is related to research activities.

Further, there is a need to clarify what are other consequences of choosing legal basis which does not benefit from OSS, including just for one international project. Does this means that DPO shall be notified and data breaches are to be reported to all MS's DPAs separately and this just for the sake of one project? What about transversal breaches, which will touch to several activities relying on different basis? Shall the same breach be reported through OSS and in parallel to all other MSs?

In case of the use of consent as legal basis, more clarification is needed in relation to the interpretation of the Article 14.2.d which states that in case of withdrawal, "the lawfulness of processing based on consent before its withdrawal" is not affected. However, in research some processing activities cannot stop, such as safety reporting or archiving for several years after the end of the study. These activities can be based on a different legal basis. However, if processing of data in a clinical research is based on the consent and patients withdraw consent at the late stage of data maturation, can data be still included in the final analysis? Or, can the results of the interim analysis still be redone to verify the correctness of it using the same dataset if some patients withdrew consent thereafter. Can this dataset be used for some types of the secondary use based on a different legal basis? For example, in the scope of inspection or audit, which is a different processing activity?

Taking an example of a research project which will collect datasets from different EU countries and provided the current heterogeneous interpretation, we may face situations where national subsets would be processed based of different legal basis within the same research project by the same data controller. In this case, not only it became



more and more challenging to conduct international studies, but it creates inequalities between patients from different countries in the way their rights are executed. Indeed, different legal basis comes with different application of rights for the data subjects.

A significant gap in the current Guidelines is the assumption that re-contacting individuals is generally feasible and that consent, whether broad or dynamic, can realistically function as a primary legal basis for scientific research across the full lifecycle of long-term, population-level data infrastructures. This assumption does not reflect the operational realities, where clinical research, disease registries, and long-term cohort omics studies often span decades, involve multiple generations of data subjects, and rely on datasets collected under statutory mandates or through heterogeneous historical pathways. In these contexts, the possibility of re-contacting individuals is frequently limited not by lack of effort, but by structural constraints: participants may be deceased, lost to follow-up, untraceable due to outdated contact details, or never directly contactable in the first place because the data originated from administrative systems, hospital records, or national registries rather than direct enrolment.

Secondary use of data is often essential for scientific progress, yet re-consent is not feasible in many cases: patients are either lost to follow-up, dead or purely not interested, and insisting on re-contact can render further research impossible or not feasible. This is particularly true for public health bodies operating under statutory mandates, where the purpose of processing is defined in law and where the GDPR explicitly provides alternative legal bases, notably Article 6(1) (e) in conjunction with Articles 9(2)(i) and 9(2)(j). The Guidelines, however, do not sufficiently acknowledge these structural limitations, nor do they provide a realistic framework for how controllers should proceed when re-contact is impossible or disproportionate. As a result, consent is implicitly framed as a default expectation, even in settings where the GDPR itself anticipates that public interest or scientific research legal bases will be more appropriate.

This lack of clarity has practical consequences. Ethics Committees, already struggling with the blurred boundary between ethical and data-protection oversight, often interpret the Guidelines as requiring re-consent for any new research purpose, even when this contradicts the GDPR, the Helsinki Declaration, or national research laws. Controllers are left navigating inconsistent expectations across Member States, with some authorities insisting on consent-based models and others recognising the legitimacy of public-interest or legitimate interest legal bases. The result is fragmentation, delays, and legal uncertainty, precisely the challenges that in many instances block the research. Moreover, the Guidelines' emphasis on dynamic consent does not account for the fact that such models presuppose stable digital infrastructures, ongoing participant engagement, and the ability to maintain up-to-date contact information, conditions that are rarely met in large-scale, population-level research systems.

***The outcome is sustained fragmentation, procedural delays, and significant legal uncertainty, which in practice operate as direct barriers to the conduct of research.***

#### **Need for clearer articulation in the Guidelines**

The Guidelines recognise broad and dynamic consent, but still assume that re-contact is generally feasible and that consent can realistically be the main legal basis in long-term, population-level infrastructures.

What is missing is an explicit recognition that consent-based models **cannot serve as the backbone** of long-term scientific research in omics technologies. The GDPR already provides a coherent legal architecture for these contexts, grounded in public interest and scientific research provisions, combined with Article 89 safeguards. The Guidelines

should reflect this reality by acknowledging that the impossibility or disproportionality of re-contact is not an exceptional circumstance but a structural feature of many research ecosystems. Without such clarification, controllers remain exposed to inconsistent interpretations, and essential research risks being hindered by expectations that are neither legally required nor operationally feasible. In health research or data-linking research, re-contact is often *impossible, disproportionate, or ethically undesirable* (e.g. deceased persons, decades-old data, large national registries).

To summarise, the guideline does not reflect the realities of personalised medicine (omics in particular) research, where:

- recontacting individuals is “often impossible or disproportionate”.
- consent is “fragile for long-term, evolving research”.
- public-interest legal bases (Art. 9(2)(j)) are more appropriate than consent.

### Concrete proposals

- Explicitly **recognise the impossibility** of recontact as a justification for relying on public interest legal bases.
- Clarify that broad consent is not required when processing is grounded in public interest.
- **Provide examples of acceptable governance** models for dynamic consent when recontact is feasible only for a subset of participants.
- Encourage Member States to **avoid imposing consent-based** regimes for genetic data where this would undermine research feasibility.
- **Explicitly harmonise structural limits of re-contact:** add a subsection stating that in:
  - long-term omics research,
  - national disease registries,
  - long-term cohorts with loss to follow-up, re-contacting individuals for each new study is often *disproportionate* under Article 14(5)(b) and Recital 62, and that reliance on **Article 6(1) (e-f) + 9(2)(i)/(j)** is normally more appropriate than consent.
- **Clarify that broad consent is not an operational precondition for omics research**, and that where Union or Member State law provides a public-interest mandate and the conditions of Article 9(2)(i) or 9(2)(j) are met, the absence of consent or the impossibility of re-contact does not prevent lawful secondary use of omics data, provided that Article 89(1) safeguards are implemented.
- **Dynamic consent as an optional enhancement, not baseline expectation:** clarify that dynamic consent is a *good practice* where digital infrastructures and stable contact are realistic (e.g. clinical trials, small cohorts), but should not be framed as a default expectation for population-level registries or surveillance systems.
- **Participation in research by design requires the processing of personal data, and introducing two separate tick boxes (one for study participation and one for GDPR consent) creates a high risk of inconsistent choices.** A patient may agree to participate in the study but inadvertently decline data processing, making full participation impossible. This approach is unreliable, confusing, and contrary to a patient-centred, accessible consent process. To avoid this contradiction, the Guidelines should explicitly allow a single, integrated consent statement that reflects the operational reality of research. A clear and patient-friendly formulation could be: *“If you agree to participate in this study, you also accept that your data will be*



*processed for the purposes of this research.*” This preserves transparency, respects autonomy, and avoids unnecessary administrative complexity, while ensuring that participants are not misled by artificial distinctions that do not reflect how research is conducted in practice.

#### 4. Ethical consent vs GDPR consent

##### Issue statement

A persistent and deeply consequential problem in the governance of clinical research is the ongoing confusion between the remit of Ethics Committees and the responsibilities assigned under the GDPR. Ethics Committees (ECs) frequently step into areas that are not legally theirs to determine, particularly regarding the choice of legal basis for processing personal data and the conditions under which secondary use may occur. Since the entry into force of the GDPR, many ECs have begun to impose GDPR consent as the legal basis for research, including in situations where this contradicts the recommendations of the EDPB, national data-protection authorities, or even the practical realities of long-term research. Ethics impose the legal basis (usually consent) for processing personal data in scope of research and in particular, the requested re-consent of the patient in case of secondary use, even though the GDPR assigns this responsibility exclusively to the controller. This tendency is not malicious; it is a symptom of a regulatory vacuum. Because the interface between ethical oversight and data-protection compliance has never been clearly articulated at EU level, ECs are left to fill the gap, often by default rather than by design.

In many instances, and often by necessity due to the absence of clear national guidance, Ethics Committees end up determining the practical legal basis for research activities, even though this responsibility formally lies elsewhere. This situation reflects the persistent uncertainty surrounding the interface between ethical and data-protection frameworks.

It should be emphasised that ethical consent is distinct from GDPR consent: the former relates to participation in research under ethical standards, while the latter concerns the lawful basis for processing personal data under data-protection law. Clarifying this distinction at EU and national levels would help ensure consistent interpretation and reduce the current reliance on ethics bodies to fill regulatory gaps.

The Guidelines acknowledge ethical standards but do not yet provide sufficient clarity on how these interact with data-protection requirements. In practice, **Ethics Committees often end up determining the effective legal basis for research**, largely because national guidance is limited or absent. This situation creates inconsistency and uncertainty across Member States.

##### Need for clearer articulation in the Guidelines

To address these issues, the EDPB should provide explicit clarification of the respective roles of Ethics Committees (ECs) and data controllers. The Guidelines should affirm that ECs are responsible for ethical and scientific oversight, not for determining the legal basis for processing personal data or for imposing GDPR-specific requirements. They should clearly distinguish between ethical consent and GDPR consent, explaining that the former governs participation in research while the latter is only one possible legal basis for processing. The Guidelines should clarify that re-consent is not required for secondary use when it is impossible or disproportionate, provided that Article 89 safeguards and ethical oversight are in place. They should also state that technical and organisational measures, including the level of pseudonymisation, are determined by the controller based on a risk analysis, and that ECs may comment on ethical implications but should not impose technical requirements. Finally, the Guidelines should encourage Member States to harmonise the interaction between ECs and DPAs and support the development of a

sector-specific code of conduct for GDPR in clinical research, as proposed in the attached document. Such a code would provide a stable, predictable framework that clarifies roles, reduces fragmentation, and supports responsible secondary use of data.

To ensure coherence, the Guidelines should explicitly distinguish between:

- **GDPR consent:** a *legal bases* under Articles 6 and 9, characterised by voluntariness, revocability, and absence of imbalance; and
- **Research or ethical consent:** a *participation agreement* governed by frameworks such as the Helsinki Declaration, the Clinical Trials Regulation, and biobank laws, which may allow waivers, deferred consent, or opt-out mechanisms.

The EDPB should clarify that:

- **How to separate ethical review from data protection assessment.**
- **Ethics Committees are not competent to impose or determine GDPR legal bases;** their role concerns ethical oversight, not data-protection law.
- **Ethical consent and GDPR consent serve different purposes** and should not be conflated in templates or review processes.
- **Secondary use of data** should be permissible without re-consent when recontacting individuals is impossible or disproportionate, provided Article 89(1) safeguards and ethical oversight are in place.

***The independence of these two assessments must be explicitly reaffirmed. Their current conflation is not a minor procedural inconvenience but one of the principal structural obstacles to efficient and legally sound research governance.***

| Aspect                  | Research consent  | GDPR consent  |
|-------------------------|---|---|
| Form                    | written consent in many (but not all) areas, dated and signed (ICH-GCP)                 | any, including oral or by ticking a box or just by providing the data |
| Explicit/implicit       | opt-out allowed in some areas (in bio-banking)  | explicit (no opt-out)   |
| Existence of exemptions | yes, provided EC approval (art 32 Helsinki declaration) or deferred consent (emergency) | no (different legal basis shall be used)                              |

**Concrete proposals:** Clarify that compliance with the GDPR is distinct from compliance with ethical and sector-specific research frameworks. Ethical oversight ensures that research is scientifically and morally sound, while GDPR compliance governs the lawful processing of personal data. These two dimensions operate independently and should not be conflated.

The EDPB should therefore:

- Explicitly state that GDPR compliance does not replace or duplicate ethical review requirements.

- Encourage clear communication between data protection authorities, ethics committees, and research sponsors to prevent overlap and confusion.
- Provide guidance illustrating how both frameworks can coexist, ensuring that data processing remains lawful under the GDPR while research continues to meet ethical and scientific standards.

## 5. PUBLIC INTEREST as legal basis

### Issue statement

A key shortcoming of the current Guidelines is the limited attention given to how “**public interest**” should be defined and operationalised as a legal basis for processing personal data for scientific and public-health research. Although sometimes treated as flexible or open-ended, public interest is **not** an undefined concept. Under the GDPR, Articles 6(1) (f), 9(2)(j) and 89 require that the **existence, scope, and conditions** of a public-interest task be **laid down in Union or Member State law**. In practice, this means controllers cannot simply invoke public interest at their discretion: the legal basis must be **explicitly established in national legislation**, including the safeguards, limits, and conditions under which research may rely on it. Instruments such as the EHDS may offer a non-exhaustive list of purposes that *may* fall under public interest, but they **do not replace** the requirement for Member States to define this basis in law.

### Context and practical implications

A concrete example of predictable and harmonised operationalisation is provided by the **French Méthodologies de Référence (MRs)** issued by CNIL. These frameworks define predefined conditions under which certain categories of health-data research are presumed to be in the public interest, subject to strict requirements such as:

- clearly defined research purpose(s), legal bases (Art. 6 + Art. 9)
- data minimisation
- independent ethical oversight
- transparency obligations
- limits on reuse

This demonstrates that public interest **can** be structured in a way that avoids arbitrary interpretation while supporting responsible research.

The Guidelines currently acknowledge the role of public interest but do not sufficiently clarify how the **sector-specific GDPR application law (EHDS)** interacts with the GDPR. The EHDS should be understood as an **operational layer** that identifies categories of secondary use and establishes procedural mechanisms for access and reuse; **not** as a redefinition of the legal concept of public interest, which remains grounded in the GDPR and national law.

### Need for clearer articulation in the Guidelines

To ensure coherence and reduce fragmentation, the Guidelines should explicitly set out what Member States must specify when establishing public-interest legal bases for research. This should include:

- which stakeholders may rely on the legal basis (public bodies, accredited research institutions, private entities under defined conditions)
- criteria for determining when a research activity qualifies as being in the public interest
- requirements for assessing the relevance, necessity, and societal benefit of the research question
- mandatory safeguards and oversight mechanisms
- conditions for data sharing and reuse
- clarify what it means to be “authorized by law” to rely on Article 6(1)(e).



- avoid the current ambiguity that leaves institutions dependent on ethics committees or DPO interpretations.
- emphasize that public interest should be defined at EU level, not left to fragmented national laws.

### Clarification on the use of Article 6(1)(e) GDPR (Public Interest) in health research

The Guidelines should explicitly recall that Article 6(1)(e) GDPR (“processing is necessary for the performance of a task carried out in the public interest”) may only be relied upon when the task is *defined in Union or Member State law*. This requirement is stated directly in the legal text: “...the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.” “The basis for the processing shall be laid down by Union law or Member State law.” (*Article 6(1)(e) and Article 6(3) GDPR*)

This means that controllers may rely on public-interest research as a legal basis only when the national legal framework explicitly covers the research task or mandate. The Guidelines should therefore encourage Member States to ensure that their national laws clearly define which research tasks fall under public interest, to avoid fragmentation and legal uncertainty, and moreover also align in requirements and definitions with the other MSs.

Of note, Article 35(10) GDPR specifies that when processing is based on such a legal basis and the law already provides for a general DPIA in the context of the adoption of the legal basis, hence, the controller does not need to conduct an additional DPIA, but must comply with all conditions and safeguards laid down in that law.

### Operational limitation: absence of the one-stop-shop mechanism

A practical constraint must also be highlighted: relying on public interest as a legal basis **excludes the GDPR one-stop-shop mechanism** (Recital 128). This has significant operational implications for cross-border research and secondary use of health data, as controllers may face multiple supervisory authorities and divergent national interpretations. Clearer guidance from the EDPB would support harmonisation and reduce fragmentation across Member States.

## 6. Transparency obligations for long-term scientific research remain insufficiently operationalised

### Issue statement

A central gap in the current Guidelines concerns the practical implementation of transparency obligations for public, academic and healthcare institutions that process personal data for scientific research over long periods of time, particularly in omics research where data pass through a chain of controllers. While the Guidelines acknowledge that controllers must ensure transparency during the entire processing period and that processors may provide information on behalf of controllers, they do not offer operational models for institutions with statutory mandates, large legacy datasets, or population-level research infrastructures. This gap is especially relevant in omics, where data originate from multiple registries, laboratories, hospitals and biobanks, and where it is often impossible to trace each dataset back to the original data holder or to re-establish direct contact with individuals.

The absence of practical guidance leaves these institutions in a structurally ambiguous position. They are expected to maintain high levels of transparency for all processing activities, including secondary use, linkage and long-term retention, yet they lack a clear, endorsed model for doing so when direct communication with millions of individuals is infeasible. The Guidelines also do not clarify how entities in the chain, such as hospitals, laboratories and diagnostic

centres, may act as communication channels without being reclassified as joint controllers. In practice, these actors are often the only realistic point of contact, but uncertainty about controllership boundaries leads to inconsistent practices and unnecessary caution. As a result, transparency obligations risk being under-implemented or implemented in ways that create legal uncertainty for all actors in the chain of controllership, particularly in omics research where data flows are complex and multi-layered.

### Need for clearer articulation in the Guidelines

The Guidelines acknowledge that processors may provide information on behalf of controllers and that exceptions under Article 14(5) exist, but they do not give *operational models* for different stakeholders in the chain of controllership and when they take data from other institutions without being the primary point of contact for the data subjects, eg. legacy datasets.

*“If controllers process personal data for long periods of time for scientific research purposes, then they should adopt appropriate measures to ensure transparency during the entire processing period.”*

The guideline does not provide operational guidance for transparency when:

- cohorts are small, dispersed, and recontact is impossible (“rare disease cohorts are typically small... recontacting individuals is often impossible”).
- data reuse spans decades.
- new findings emerge (e.g., reinterpretation of variants).

### Concrete proposals

- **Model “layered transparency” for research community:** recommend a layered approach for institutions:
  - **Primary layer:** institution, research level, public level, continuously updated web portal describing all main research programmes, data sources, safeguards, and rights.
  - **Secondary layer:** registry- or study-specific pages, with lay summaries and contact points.
  - **Tertiary layer:** on-demand detailed information via central contact (DPO/research office).
- **Clarify acceptable reliance on Article 14(5)(b):** provide criteria for when individual notification is “impossible or would involve a disproportionate effort” in population research, e.g.:
  - very large cohorts/registries,
  - outdated contact details,
  - linkage of multiple administrative sources, combined with robust public communication and easily accessible information.
- **Processors as communication channels:** confirm that hospitals, labs, or other processors may:
  - provide information at point of care,
  - handle access/erasure requests on behalf of the controller, under documented instructions, without becoming joint controllers.

## 7. Proportionality and legitimate interest assessment (LIA)

### Issue statement

A significant gap in the current Guidelines is the absence of a structured, research-tailored approach for applying legitimate interest as a legal basis for scientific processing. Although the Guidelines acknowledge that genuine scientific research carries “significant weight” in a legitimate interest assessment, they stop short of providing a

methodological framework that reflects the complexities of research practice. This omission is particularly striking given the challenges in the sector of omics research where controllers already face fragmented interpretations of legal bases, inconsistent expectations from Ethics Committees, and uncertainty about how to justify the necessity and proportionality of processing in long-term, multi-stakeholder research environments.

In clinical and public health research, controllers routinely process sensitive data for purposes that evolve, require linkage across datasets, and depend on maintaining identifiable or pseudonymised information to ensure scientific validity, minimise bias, and enable follow-up. Yet the GDPR's general legitimate Interest Assessment (LIA) model was designed for commercial contexts, not for research ecosystems governed by ethical oversight, statutory mandates, and Article 89 safeguards. Without a research-specific LIA structure, controllers are left to adapt generic templates that do not capture the scientific rationale, methodological constraints, or public-interest dimension inherent to research. This lack of tailored guidance increases the risk of inconsistent assessments across Member States, reinforces the already-documented fragmentation in GDPR interpretation, and leaves controllers exposed to challenges from Ethics Committees or DPAs who may not share a common understanding of how legitimate interest applies in research.

Research often operates within hybrid governance models, involving public bodies, academic institutions, private partners, and international collaborators, where the choice between Article 6(1) (e) and 6(1)(f) is not always straightforward.

In the absence of clear criteria, controllers may default to consent even when it is inappropriate or unworkable, or they may avoid legitimate interest altogether due to fear of inconsistent scrutiny. This dynamic exacerbates the legal uncertainty already affecting secondary use, long-term follow-up, and methodological innovation. A research-specific LIA framework would not only support controllers in documenting their reasoning but would also promote harmonisation, reduce unnecessary reliance on consent, and provide regulators with a consistent basis for evaluating research activities.

The guideline does not explain how proportionality should be assessed when:

- anonymisation destroys scientific value (“generalisation, suppression, aggregation tend to destroy data utility”).
- genetic data inherently carries identifiability risk (“uniquely identifying... shared with biological relatives”).
- risk depends on data type (WGS vs SNP vs methylation).

### **Need for clearer articulation in the Guidelines**

The Guidelines would benefit from a clearer and more operational articulation of how Article 6 and Article 9 GDPR must be combined when processing sensitive data for scientific research. As currently drafted, the document places disproportionate emphasis on consent-based models, including broad and dynamic consent, even though the GDPR itself recognises that consent is rarely a feasible or reliable legal basis in research involving vulnerable populations, long-term follow-up, multi-stakeholder infrastructures, or secondary use. This approach risks reinforcing the very fragmentation the Guidelines seek to remedy. In practice, research sponsors and infrastructures need legal bases that are stable, scalable, and compatible with cross-border governance. The legitimate interest under Article 6(1)(f), combined with Article 9(2)(j) for scientific research, is often the only workable basis that both preserves the one-stop-shop mechanism and reflects the real operational conditions of research. Legitimate interest is explicitly recognised in the Guidelines as a valid ground for scientific research, and the EDPB acknowledges that genuine

research carries significant societal weight in the balancing test. The Guidelines should therefore clarify when and how controllers may rely on this combination, and provide concrete criteria for distinguishing research-motivated legitimate interests from commercial product development. Such clarification would support harmonisation, reduce unnecessary reliance on consent, and offer a proportionate, risk-based pathway that aligns with the GDPR's architecture and with the needs of modern research ecosystems.

The Guidelines acknowledge that scientific research can be a legitimate interest and that "significant weight" can be given to it, but they do not provide a *research-specific LIA framework*.

What is missing, therefore, is a structured, operational model that translates the legitimate interest test into the realities of scientific research, one that recognises scientific validity, public value, methodological necessity, and the role of safeguards as integral components of the assessment. Without such a framework, legitimate interest remains theoretically available but practically under-used, contributing to the broader uncertainty that continues to hinder lawful, high-quality research across the Union.

### Concrete proposals

- **Clarification on the roles and responsibilities:** the responsibility for conducting and documenting the Legitimate Interest Assessment (LIA) must remain with the researcher or sponsor acting as the data controller. Regulatory bodies and ethics committees are responsible for evaluating the overall research protocol, including its ethical soundness and the protection of participants' rights, but they should not prescribe or impose how the LIA itself is performed.
- **Understanding and documenting the reasonable expectations of data subjects:** this is a core element of the balancing test in a Legitimate Interest Assessment. In scientific and clinical research, this assessment must reflect the **specific context of data collection**, the **nature of the relationship between participant and controller**, and the **foreseeability of secondary use**. This step is often poorly understood and inconsistently applied, leading to unnecessary re-consent demands, misinterpretation by ethics committees, and fragmentation across Member States. Specifically, a research-specific LIA must include a clear and well-reasoned assessment of the reasonable expectations of data subjects, grounded in the context in which their data were originally collected. Controllers should document the nature of the initial interaction with participants, the information provided at the time (including any reference to future scientific use), and the extent to which further processing within the same disease area, research programme, or public health mandate would have been foreseeable. This assessment should also consider the time elapsed since collection, changes in legal or technological context, and whether re-contact is feasible or appropriate in long-term or population-level research. Expectations must be evaluated in light of the relationship between participant and controller: clinical trial sponsors, healthcare providers, registries, and public health authorities each create different baselines of trust and foreseeability. Controllers should capture this reasoning in the LIA, DPIA, and Record of Processing Activities, demonstrating how transparency materials, ethical oversight, pseudonymisation, and secure environments ensure that secondary use aligns with what individuals can reasonably anticipate in a regulated research setting.
- **Require risk-tiering** of genetic/omic data (identity, attribute, group disclosure risks).
- **Provide proportionality** examples for each tier (e.g., WGS always high-risk; SNP panels context-dependent).
- **Clarify that proportionality** must balance scientific necessity and privacy risk, not assume anonymisation is always possible.

- **Encourage use of secure processing environments** as a proportionality safeguard.
- **Research-tailored LIA template:** suggest that for Article 6(1)(f) in research, controllers explicitly assess:
  - **Interest:** scientific validity, public value, alignment with recognised research integrity standards (e.g. <https://allea.org/code-of-conduct/>).
  - **Necessity:** why identifiable/pseudonymised data are needed (e.g., dataset linkage, follow-up, bias control).
  - **Balancing:** risk factors (vulnerability, stigma, data sensitivity) vs safeguards (pseudonymisation, secure environments, ethics oversight, PETs).
- **Proportionality as cross-cutting principle:** encourage explicit documentation of:
  - data minimisation choices (variables, sample size),
  - access controls (role-based, time-limited),
  - publication safeguards (aggregation, suppression).
- **Clarify when LIA is *not* appropriate:** indicate that for statutory public health tasks, Article 6(1)€ is normally more appropriate than 6(1)(f), while LIA may be suitable for:
  - public-private partnerships,
  - method development,
  - exploratory analyses outside a strict legal mandate.

## 8. Lack of operational clarity on documenting and applying the presumption of compatibility for scientific research

### Issue statement

In our understanding, GDPR provides three key mechanisms for further processing:

- i) Consent
- ii) Union and/or National law (which itself can provide legal basis for further processing as per Recital 50 GDPR)
- iii) Compatibility of purposes

Given that re-consenting is not always feasible in clinical research context, due to the -mentioned reasons, compatibility of purposes can serve as another option. This, in our view, implies that the controller, in order to enable further processing, is required to justify the compatibility of the initial and secondary purposes of processing. However, this aspect still raises debate in the view of the statement of the Article 5.1.b as some interpretations suggest it releases data controllers from the need to perform the compatibility test. This point of view is supported by the Recital 50 GDPR which states that in case of successful compatibility test, separate legal basis is not required, but at the same time clarifies that “further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations”. Clarification of this element would help sponsors to understand applicable requirements.

Last but not least, the Recital 50 GDPR may suggest that the need for compatibility test may differ from the legal basis used initially (not required in case of the initial consent followed by important objectives of general public interest versus required in case of legitimate interests initially used).

**A link between initial and secondary purposes of processing:** the term ‘link between purposes may have different interpretations and seems to be quite broad and vague. From our point of view, establishing a link between initial and secondary purposes can be a major indication of their compatibility and thus the main criterion in the

compatibility methodology. The more purposes are related and connected to each other, the bigger chance they are compatible, regardless four following criteria. This topic is in relation to the previous considerations about the granularity of purposes. The question that arises here is how different the purposes should be, in order to measure the “distance” between them. In the research environment, it is not always possible to predict further projects of research and thus the need for further processing. For instance, at the stage of data collection none can anticipate what exactly tests/analyses will be performed on patient’s data. Hence, “encapsulated”, broader purpose of processing should be allowed in our view.

**The context of data collection:** the question to ask here is whether the further processing will take place in the same context as the data collection. We believe that providing illustrative examples of contexts to be considered would be helpful to the research community. By context we understand, for example, the use of data for commercial purposes versus non-commercial purposes. Again, transparency is the key and researchers shall be able to inform data subjects from the onset about all possible processing activities that can be made, eventually providing some choices where relevant. Narrowing the scope of information provided to data subjects from the onset is not transparent and is detrimental to both data subjects and to the research and ultimately the EU capacity of innovation.

**The nature of personal data:** healthcare research always use special categories of data. However, it may also use genetic data. Currently many countries limit the possibilities of use of genetic data considerably and in a very divergent way (based on the article 9.4); this without making any distinction between the type of the data at stake (somatic mutations versus germ line mutations, primary DNA sequence versus results of analysis of the status of a single frequent mutation etc...). This approach is detrimental to research and does not always correspond to any higher protection of data subjects. In our view, there is a need to apply a risk-based approach and all genetic data shall not be treated the same, some do not present more risk as compared to any other health data, others do. On the other hand, data as rich as the primary sequence, whereby less than hundred pairs of nucleotides make it unique, are not eye readable to make any sense of them. “Reading” genetic data requires highly sophisticated specialised software, which in a way protects them.

**Consequences of the intended further processing (risk assessment):** there should be a balance made between negative and positive or rather beneficial consequences for data subjects. Examples of positive consequences can be development of effective and acceptably safe medicines or a sense of contribution to innovation aiming to improve either the condition or the knowledge about it. In a way this is a type of the risk benefit assessment well known in the field of research involving humans. In our view this is where the contribution of an ethics committee would be of a value.

**Appropriate safeguards in both initial and further processing:** safeguards used for the initial processing should be re-evaluated if they are relevant for further processing. Such safeguards as encryption and pseudonymization require extra efforts, due to constant technological developments. In many, but not all cases double pseudonymisation can be applied. However, beyond its benefits for the protection of confidentiality, it involves risks to the data accuracy and thus its implementation cannot be automatic but needs to be considered case by case considering all, including technical aspects.

While the Guidelines correctly recall that further processing for scientific research is presumed compatible with the initial purpose under Article 5(1)(b), they stop short of explaining how controllers should document and rely on this

presumption in complex, multi-purpose research environments. In practice, controllers face uncertainty when secondary use involves layered datasets, mixed legal bases, or evolving research aims, for example, when registry data collected for surveillance are later used to study treatment outcomes or genetic correlations. The Guidelines' current wording leaves open whether such reuse requires a separate compatibility test or how controllers should record their reasoning to demonstrate compliance.

This gap is particularly problematic for public health and biomedical research, where data are collected under statutory mandates and reused over decades. Lawful initial collection, genuine scientific purpose, and implementation of Article 89(1) safeguards should suffice to rely on the presumption of compatibility. Yet without explicit guidance, controllers, ethics committees, and data-access bodies interpret this differently, leading to fragmentation and unnecessary procedural burdens. The absence of examples, such as cancer registries used for survival analyses or vaccine surveillance data reused for effectiveness studies, further limits the practical value of the presumption.

What is missing is a clear, operational framework that tells controllers how to document compatibility reasoning within their records of processing and DPIAs, and how to apply the presumption consistently across multi-purpose research infrastructures. Explicit criteria and real-world illustrations would ensure that the presumption functions as intended: to enable lawful, transparent, and proportionate secondary use of data for scientific and public-interest research, without redundant compatibility testing or conflicting national interpretations.

#### Need for clearer articulation in the Guidelines

The Guidelines rightly recall the presumption of compatibility for further processing for scientific research, but they do not give enough *practical guidance* on how controllers should document and rely on it in complex, multi-purpose environments.

#### Concrete proposals

- **Clarify scope of the presumption:** state explicitly that the presumption applies when:
  - the initial collection was lawful,
  - further use is genuinely scientific research (per the six key-indicative factors),
  - Article 89(1) safeguards are implemented.
- **No separate compatibility test, but documented reasoning:** recommend that controllers:
  - record in their RoPA/DPIA that a given secondary use falls under Article 5(1)(b) second sentence,
  - briefly justify why it remains within the same broad research/public health objective.
- **Examples for public health:** add examples such as:
  - using cancer registry data collected for incidence monitoring to study treatment patterns and survival,
  - using surveillance data to evaluate vaccine effectiveness, as *presumed compatible* further processing.

### 9. Lack of guidance on proportionate, risk-based pseudonymisation in research

#### Issue statement

The Guidelines do not address a critical operational gap: how controllers should determine proportionate pseudonymisation measures in scientific and clinical research. In practice, overly rigid interpretations, such as

mandatory double pseudonymisation, are increasingly imposed by some authorities or ethics bodies, even though such requirements may be scientifically unnecessary, operationally burdensome, or even harmful to sample integrity, particularly for frozen or fragile biological materials. Evidence shows that double pseudonymisation can require additional manipulation of samples, increasing the risk of degradation and compromising the scientific validity of downstream analyses.

What is missing from the Guidelines is explicit recognition that pseudonymisation is not a one-size-fits-all obligation. Controllers must retain the discretion to select safeguards that are proportionate to the actual risks of the processing, taking into account the nature of the data, the research design, the feasibility of re-identification, and the Article 89(1) safeguards already in place. Without clear guidance, controllers face inconsistent expectations across Member States, and research infrastructures are exposed to unnecessary procedural barriers that do not enhance data protection.

The Guidelines should therefore clarify that the choice between single pseudonymisation, double pseudonymisation, or alternative safeguards must be based on a documented, risk-based assessment by the controller, who is best placed to evaluate scientific needs, technical constraints, and the risk profile of the processing.

A further gap with significant consequences for omics innovation, personalised medicine, and rare-disease research is the absence of guidance on how proportionate pseudonymisation should be determined in practice. Omics datasets, whether whole-genome sequences, transcriptomes, methylomes, or multi-omics profiles, are inherently identifying, but they also require high-fidelity, minimally manipulated biological material to remain scientifically valid. Rigid interpretations emerging in some Member States, such as mandatory double pseudonymisation for all omics workflows, ignore the scientific and technical realities of these data types. For rare-disease cohorts and personalised-medicine pipelines, additional manipulation of frozen or fragile samples can degrade nucleic acids, compromise variant calling, and undermine downstream analyses. Over-engineering pseudonymisation not only fails to meaningfully reduce re-identification risk, given the intrinsic identifiability of omics data, but also creates barriers to diagnosis, delays in treatment optimisation, and unnecessary fragmentation of research infrastructures. The Guidelines should therefore explicitly recognise that pseudonymisation in omics research cannot be treated as a uniform procedural requirement. Controllers must retain the ability to select safeguards that are proportionate to the actual risk profile, the scientific purpose, and the Article 89(1) framework already in place. Without such clarification, Europe risks imposing procedural burdens that do not enhance data protection but do directly impede the development of personalised medicine, slow rare-disease diagnostics, and weaken the competitiveness of EU omics innovation.

#### **Need for clearer articulation in the Guidelines**

The Guidelines should explicitly recognise that controllers must retain the discretion to determine the appropriate level of pseudonymisation and related safeguards based on a documented, risk-based assessment. Overly prescriptive requirements, such as mandating double pseudonymisation in all cases, do not reflect the operational realities of clinical and translational research and may even compromise sample integrity, particularly for frozen or fragile biological materials. The decision on whether single or double pseudonymisation, or alternative safeguards, is proportionate should rest with the controller, who is best placed to evaluate scientific needs, technical constraints, and the risk profile of the processing. The Guidelines should therefore clarify that controllers may select proportionate measures tailored to the specific research context, provided that the reasoning is documented and

that Article 89(1) safeguards are implemented, allowing controllers to choose proportionate safeguards based on risk.

EDPB could acknowledge the practical impossibility of anonymisation in most research contexts:

- Move away from theoretical debates on anonymisation.
- Emphasise that true anonymisation is rarely possible in research.
- Focus instead on risk-based safeguards and secure processing environments.

### Concrete proposals

#### 1. Clarify controller discretion and risk-based pseudonymisation (Section 8.3 / 8.5)

Add the following paragraph at the end of Section 8.3 (“Anonymisation and pseudonymisation of personal data”) or, alternatively, as a new chapeau paragraph in Section 8.5 (“Determination of other types of appropriate safeguards”):

*“Pseudonymisation is not a one-size-fits-all obligation. In line with the risk-based approach and the principle of accountability, controllers should retain the discretion to determine the appropriate level and modality of pseudonymisation (e.g. single pseudonymisation, double pseudonymisation, or alternative safeguards) on the basis of a documented risk assessment. This assessment should take into account, inter alia, the nature and sensitivity of the data, the research design, the feasibility of re-identification in the concrete context, and the safeguards implemented pursuant to Article 89(1) GDPR. Overly prescriptive requirements, such as mandating double pseudonymisation in all cases—may be neither necessary nor proportionate and, in some settings (e.g. manipulation of frozen or otherwise fragile biological samples), may even compromise sample integrity and thereby the scientific validity of the research.”*

#### 2. Clarify the role of ethics / information security bodies versus the controller (Section 8.2 / 8.5)

Insert the following paragraph in Section 8.2 (“Risk analysis and data protection impact assessment”) or Section 8.5:

*“Ethics committees, information security committees and similar oversight bodies play a crucial role in reviewing the overall research protocol, including the protection of participants’ rights and interests. However, the decision on which technical and organisational measures (TOMs) are implemented, such as the specific form of pseudonymisation or data linkage architecture, remains the responsibility of the controller, who must conduct the DPIA, assess residual risks and document the chosen safeguards. Oversight bodies should therefore avoid prescribing specific TOMs as a matter of principle and instead focus on advising and collaborating with controllers to identify feasible, proportionate solutions that adequately mitigate risks while preserving scientific robustness.”*

#### 3. Provide guidance on data linkage in research infrastructures (Section 2.2 / 8.3)

Add the following paragraph in Section 2.2 (“Research data infrastructures”) or Section 8.3:

*“In many scientific research settings, especially in longitudinal and registry-based studies, controlled data linkage is essential to ensure data quality, follow-up and clinically meaningful outcomes. The Guidelines should be understood as permitting pseudonymisation architectures that enable secure linkage (e.g. via trusted third parties or key-management services), provided that: (i) identifiers are kept logically and organisationally separate from research datasets; (ii) access to linkage keys is strictly role-based and auditable; and (iii) the linkage design is described in the DPIA and forms part of the Article 89(1) safeguards. Controllers should document why the chosen linkage model is necessary for the research aims and how the residual re-identification risk is controlled.”*

## 10. International data transfers for scientific research are insufficiently addressed

### Issue statement

The Guidelines provide almost no practical guidance on international data transfers, even though cross-border collaboration is structurally necessary for scientific, clinical, and public-health research. Existing GDPR transfer tools, such as SCCs, TIAs, ad hoc clauses, and Schrems II supplementary measures, were designed for bilateral commercial service provision, not for complex research ecosystems involving mixed controller/processor roles, federated infrastructures, public bodies, and cloud-based analytics.

The Guidelines do not address how GDPR transfer rules apply when research involves mixed C→C / C→P / P→C flows, nor how to manage transfers to public research bodies that cannot sign SCCs in commercial form (e.g., ministries, national institutes, public health authorities). They also do not clarify whether cloud hosting or remote access by non-EU providers constitutes a “transfer” in research settings, nor how to apply Schrems II supplementary measures when data are pseudonymised, held in secure environments, or accessed only through controlled queries.

The absence of research-specific guidance leaves controllers facing inconsistent national interpretations, unnecessary delays, and legal uncertainty, particularly for secondary use, registries, surveillance systems, and long-term cohort infrastructures. This gap is especially problematic given the Guidelines’ own recognition that scientific research is a fundamental EU objective and that data disclosure to secondary users may itself constitute a research purpose. Without a nuanced interpretation of transfer rules, the presumption of compatibility and Article 89(1) safeguards cannot function as intended.

### Need for clearer articulation in the Guidelines

The Guidelines should explicitly acknowledge that international transfers in scientific research differ fundamentally from commercial outsourcing. Research infrastructures often involve :

- multi-party consortia with evolving roles over time,
- public bodies bound by statutory secrecy and ethics regimes,
- federated or distributed analytics where data do not leave the EEA,
- secure processing environments where only aggregated outputs are exported,
- pseudonymisation of datasets with no direct identifiers,
- data-visiting models where algorithms travel to the data, not vice-versa.

These models do not fit neatly into the SCC-based transfer architecture. The Guidelines should therefore articulate how controllers can comply with GDPR Chapter V (Transfers of personal data to 3<sup>rd</sup> countries or international organisations) in a way that preserves scientific integrity, avoids unnecessary re-engineering of infrastructures, and recognises the safeguards already embedded in research governance (ethics review, professional secrecy, Article 89(1) measures, processing secure environments, reliance on PET). Without such clarification, controllers are left to interpret transfer rules in ways that may be overly restrictive, inconsistent across Member States, or incompatible with the operational realities of public-health and clinical research.

## Concrete proposals

1. **Call for an EDPB-endorsed research Code of Conduct with:**
  - model clauses for mixed controller/processor roles;
  - rules for transfers to public research bodies unable to sign SCCs;
  - safeguards for cloud-based analytics;
  - risk-based transfer assessment tailored to scientific research.
2. **Clarify how Art. 49 derogations may be used for rare-disease research when no stable transfer tool exists.**
3. **Provide examples of acceptable governance for federated analysis.**
4. **Support development of a sector-specific Code of Conduct on international transfers (Article 40 GDPR)**

Add a new paragraph in Section 8.5 (“Determination of other types of appropriate safeguards”):

*“Given the structural importance of cross-border collaboration in scientific research, the EDPB encourages the development of a sector-specific Code of Conduct under Article 40 GDPR addressing international transfers in research. Such a Code could include: (i) model clauses for mixed controller/processor constellations typical of research consortia and data repositories; (ii) conditions for transfers to public research bodies unable to sign SCCs in commercial form; and (iii) safeguards for federated and cloud-based analytics, including data-visiting and secure processing environments.”*

5. **Provide research-tailored guidance on Transfer Impact Assessments (TIAs) while recognising the burden and propose workable alternatives for research: sector-specific flexibility or alternative mechanisms.**

Insert a new subsection under Section 8.2 (“Risk analysis and DPIA”):

*“When conducting Transfer Impact Assessments in scientific research, controllers may take into account contextual safeguards such as: (i) pseudonymisation with separation of keys; (ii) use of secure processing environments where only query results leave the EEA; (iii) the existence of statutory professional secrecy, ethics oversight, or research-specific confidentiality obligations applicable to foreign public bodies; and (iv) the scientific necessity of maintaining data integrity for longitudinal or cohort-based analyses. The Guidelines should provide examples illustrating how these factors can reduce residual risk in a Schrems II assessment.”*

6. **Clarify the application of Schrems II supplementary measures in research settings**

Add to Section 8.3 (“Anonymisation”):

*“In research contexts, supplementary measures may include: (i) pseudonymisation where re-identification keys remain exclusively within the EEA; (ii) controlled remote access through secure*



*environments; (iii) output-checking mechanisms; and (iv) governance-based safeguards such as ethics approval and professional secrecy. The Guidelines should clarify that these measures may be sufficient where the data importer cannot access identifiable data or where only aggregated results are exported.”*

**7. Clarify when cloud hosting by non-EU providers constitutes a “transfer”**

Add to Section 3.1.1 (“Presumption of purpose compatibility”) or Section 8.3:

*“The Guidelines should clarify whether cloud hosting or remote maintenance by non-EU providers constitutes a transfer in research settings, particularly where: (i) data are encrypted or pseudonymised; (ii) keys remain exclusively within the EEA; and (iii) the provider has no technical ability to access intelligible data. Clear criteria are needed to avoid divergent national interpretations.”*

**8. Provide guidance for transfers to academic, hospital, research institutions and public bodies unable to sign SCCs**

Add to Section 4.4.3 (“Derogations under Union or MS law”):

*“The Guidelines should address transfers to public, academic research bodies (e.g., university hospitals, ministries, national institutes, public health authorities) that cannot sign SCCs in commercial form. Controllers should be able to rely on alternative safeguards, such as statutory secrecy, ethics oversight, or binding administrative arrangements, where these provide protection essentially equivalent to SCCs.”*

**9. Clarify the role of Article 49(1)(d) public-interest derogation for urgent research**

Add to Section 4.4.3 or Section 6.3:

*“The Guidelines should clarify when Article 49(1)(d) (‘important reasons of public interest’) may be used for urgent cross-border public-health research, such as outbreak investigations or rapid response studies. Use of this derogation should be embedded in documented governance and not serve as a routine substitute for structural transfer mechanisms.”*

## 11. Genetics, omics, special categories of data and research infrastructures

### Issue statement

The current Guidelines treat genetic and omics data under the broad umbrella of “special categories of personal data,” but they do not reflect the heterogeneous risk spectrum within omics datasets nor the practical impossibility of anonymising genomic data in many research contexts. They also overlook the distinct governance needs of rare-disease and personalised-medicine (gen)omics, where identifiability is structurally higher and long-term reuse is scientifically indispensable. As a result, controllers lack operational guidance on how to calibrate safeguards, retention, and reuse for different types of omics data, and ethics committees often impose uniform requirements that do not align with GDPR’s risk-based approach. To address these gaps, it is essential to recognise that research infrastructures are not ancillary to research but constitutive of it: they provide the secure, quality-assured, expertise-driven environments through which omics data are generated, curated, accessed, and reused.

Separating infrastructures from research, fragmenting roles, or treating interdependent components of the omics ecosystem as isolated “purposes” would undermine scientific coherence and weaken safeguards. Excessive granularity that ignores how infrastructures, researchers, and governance mechanisms interlink creates structural fragility rather than protection. A sustainable model must therefore remain value-driven and safeguard-oriented, recognising the cumulative impact of the ecosystem as a whole rather than enforcing siloed interpretations that harm research, data quality, and public benefit.

### Need for clearer articulation in the Guidelines

A research-specific GDPR Good Practice framework, similar to ICH-GCP<sup>7</sup>. The Guidelines should explicitly recognise that:

**1. Genetic/omics data are not homogeneous:** WGS/WES, multi-omics datasets, tumour-only somatic profiles, and single biomarkers present very different identifiability, disclosure, and linkage risks. Treating them as a single category leads to disproportionate safeguards and unnecessary barriers to research.

**2. Full anonymisation of omics data is rarely feasible:** the Guidelines currently imply anonymisation as a preferred safeguard, but do not acknowledge that:

- genomic data are inherently identifying,
- anonymisation often destroys scientific value,
- pseudonymised omics data remain personal data and must be governed accordingly.

**3. Controllers need explicit support for long-term retention and re-analysis:** rare-disease and (gen)omics require decades-long retention and periodic reinterpretation. The Guidelines do not provide examples or criteria for documenting necessity under Article 5(1)€ and Article 89(1).

**4. Ethics committees often impose technical measures:** the Guidelines do not clarify that the controller, not ethics bodies, is responsible for determining proportionate safeguards through DPIA and risk assessment. This gap leads to rigid requirements (e.g., mandatory double pseudonymisation) that may harm sample integrity or impede data linkage.

### Concrete proposals

1. **Provide a taxonomy of genetic/omic data (raw vs derived; inherited vs acquired).**
2. **Clarify when epigenetic data fall under Art. 4(13).**
3. **Encourage Member States to avoid divergent definitions that undermine cross-border research.**
4. **Provide examples of risk-appropriate safeguards for each data type (e.g., WGS → SPE only; variant lists → controlled access).**

---

<sup>7</sup> [ICH Official web site : ICH](#)

**5. Introduce a risk-graded framework for genetic and omics data.** Add a subsection under 8.4 Safeguards when processing genetic or biometric data:

*“Controllers should distinguish between different categories of genetic and omics data when conducting DPIAs and determining safeguards. A risk-graded approach may include:*

- *High-risk: whole-genome/whole-exome sequencing, multi-omics datasets, rare-disease profiles;*
- *Intermediate-risk: tumour-only somatic sequencing, panels of common variants;*
- *Lower-risk (still special category): single biomarkers with limited identifiability.”*

This aligns with GDPR’s risk-based approach and avoids one-size-fits-all safeguards.

**6. Why secure research infrastructures are integral to lawful omics data use.** Add to 8.3 Anonymisation and pseudonymisation:

*“The Guidelines recognise that full anonymisation of genomic and multi-omics datasets is rarely achievable in practice without eliminating scientific utility. Controllers should therefore treat pseudonymised omics data as personal data, apply Article 9(2) conditions and Article 89(1) safeguards, and avoid relying on unrealistic expectations of ‘anonymous genomics’. A practical interpretation of GDPR Article 9(2)(j) must recognise that research infrastructures which already apply GDPR safeguards and sector-specific rules are themselves part of the research ecosystem, not external or ancillary to it. Modern genomics and multi-omics research depends on secure, well-governed environments that implement pseudonymisation, controlled access, role-based permissions, auditability, data-quality standards, and expert oversight, all safeguards explicitly required under Article 89 GDPR. Infrastructures such as biobanks, federated analysis platforms, secure processing environments, and omics-data repositories therefore meet the requisites of “processing for scientific research purposes” when they operate under these conditions. This interpretation is consistent with the European Health Data Space (EHDS) Regulation, which recognises that electronic health and genomic data may be shared for secondary use only through controlled-access mechanisms. EHDS does not limit research to a single centralised structure; instead, it establishes a governance model in which multiple compliant infrastructures may provide access, provided they apply GDPR-level safeguards and sector-specific protections. It is also aligned with the WHO 2023–2024 Guidance on Human Genome Data<sup>8</sup>, which emphasises that genomic data must be handled within secure, quality-assured, expertise-driven infrastructures and warns that governance models should avoid creating barriers to legitimate scientific research. WHO explicitly frames infrastructures as integral components of the scientific process, ensuring quality, integrity, and interpretability across the entire lifecycle of genomic data. For these reasons, infrastructures that implement GDPR Article 9 safeguards, Article 89 research protections, and relevant sector-specific rules (e.g., clinical genomics, biobanking, AI-based analysis) should be considered fully within the scope of “scientific research purposes.” They enable, not merely support, the lawful and responsible use of omics data, and they ensure that data can be shared under controlled access in a manner consistent with both EHDS and international best practice. No research can take place in the absence of data. Data are the indispensable starting point of every scientific question, method, and conclusion.”*

**7. Provide explicit guidance on retention and re-use of omics data.** Add to 3.2 Storage limitation:

*“Long-term retention of omics data may be justified where future re-analysis is scientifically necessary (e.g., rare diseases, evolving interpretation of variants). Controllers should document:*

- *the scientific rationale for long-term retention;*
- *the safeguards ensuring restricted access;*
- *the governance for future re-use (e.g., ethics/ELSI review, Article 9(2)(j) conditions).”*

<sup>8</sup> <https://www.who.int/publications/i/item/9789240102149>

This supports lawful long-term storage without requiring repeated re-consent.

**8. Clarify the controller’s role in determining safeguards.** Add to 8.2 Risk analysis and DPIA:

*“The determination of appropriate technical and organisational measures, including the degree of pseudonymisation, rests with the controller, based on a documented DPIA. Ethics committees and information-security committees may provide advice, but should not impose specific technical measures unless required by Union or Member State law.”*

**9. Add guidance on data linkage for omics.** Add to 8.5 Determination of other types of appropriate safeguards:

*“Where scientific validity requires linkage of omics data with clinical, imaging, or registry data, controllers should document:*

- *the necessity of linkage for the research purpose;*
- *the safeguards preventing unauthorised re-identification;*
- *the governance for linkage keys (e.g., held by an independent trusted third party).”*

## 12. Unharmonised interpretation of complex research roles

### Issue statement

In clinical and health research, roles are structurally complex: sponsors, investigators, sites, laboratories, and data platforms may each act as controller, joint controller, or processor for different processing operations. Investigators often “wear different hats” (care vs research), and DPAs across Member States diverge on sponsor–investigator controllership models. This creates uncertainty for contracts, transparency notices, and allocation of responsibilities.

### Need for clearer articulation in the Guidelines

The Guidelines should move beyond generic controller/processor definitions and provide research-specific illustrations that reflect real-world clinical-trial and registry architectures, including the dual role of investigators and the coexistence of multiple controllers in a single project.

### Concrete proposals

- **Explicitly harmonise “multi-role” actors in research**
  - **Sponsor:** Usually a controller for research-related processing (designs protocol, defines endpoints, determines data flows).
  - **Investigator / site :**
    - controller for patient-care processing (medical record, routine care, safety follow-up beyond protocol);
    - controller or joint controller *or* processor for research-specific processing, depending on whether they co-design the protocol and essential means.
  - The Guidelines should endorse this “different hats” model and recommend that controllers *describe it explicitly* in DPIAs, contracts, and information notices.
- **Provide structured criteria for sponsor–investigator roles**
  - **Joint controllership indicators :**
    - co-drafting of protocol and essential means;

- shared decisions on endpoints, data categories, and follow-up;
- shared scientific responsibility for the study design.
- **Processor indicators :**
  - investigator follows a sponsor-defined protocol without meaningful influence on purposes and essential means;
  - processing is limited to implementing sponsor instructions for data collection and reporting.
- **Address divergent DPA interpretations with a “safe harbour” logic**
  - **Acceptable role models:** The Guidelines could present a small set of *acceptable* role configurations (e.g. “Model A: sponsor controller / investigator processor for research data; investigator separate controller for care data”; “Model B: sponsor and investigator joint controllers for research data”), and state that, if one of these models is transparently adopted and documented, it should generally be considered compliant.
  - **Documentation and transparency :**
    - recommend that sponsors/controllers keep a “controllership map” describing roles per processing operation and per jurisdiction;
    - require that this map underpins:
      - joint-controller arrangements (Art. 26),
      - data-processing agreements (Art. 28),
      - and layered information to participants (who does what, for which purpose, under which legal basis).
- **Clarify that ethics-committee views on roles are not determinative under GDPR**
  - The Guidelines should state that ethics committees may *comment* on roles, but the legal qualification and decision of mapping controller/processor’s roles is ultimately a GDPR question for controllers and, where relevant, DPAs.
  - Where ethics-committee expectations diverge from DPA guidance, controllers should:
    - document the conflict,
    - prioritise compliance with binding data-protection law,
    - and, where possible, explain to ethics committees how the chosen model still ensures high protection for participants.

### 13. Rights of data subjects whose data are used in research

#### Issue statement

The Guidelines acknowledge that certain rights, particularly erasure (Art. 17) and objection (Art. 21), may be restricted when their exercise would “render impossible or seriously impair” scientific research. However, they do not provide **operational guidance** on how these rights should be exercised in **multi-layered research ecosystems**, where:

- data flow across **controller-to-controller** and **controller-to-processor** chains,
- sponsors typically hold only **pseudonymised datasets** and **cannot re-identify individuals**,
- investigators “wear different hats” (clinical care vs. research),
- research institutions, university (research) hospitals, registries, repositories, and public-health bodies act as **independent controllers**,
- and data subjects may not know which entity is responsible for which part of the processing.

In practice, most data-subject requests are handled by the investigator or treating site, because they are the only party that can directly identify the participant. The sponsor, who receives only pseudonymised data, has neither the legal right nor the operational ability to access the re-identification key. As a result, sponsors are not expected, and should not be required, to communicate directly with clinical-trial participants to fulfil data-subject rights.

The absence of operational models creates uncertainty for controllers, ethics committees, and DPAs, and risks inconsistent implementation across Member States.

### Need for clearer articulation in the Guidelines

The guideline does not address:

- how rights apply when recontact is impossible;
- how to handle reinterpretation of genomic findings;
- how to manage rights in small rare-disease cohorts where identifiability is inherent;
- how rights interact with statutory public-health mandates.

The Guidelines should clarify **how data subject rights are to be exercised and fulfilled in complex research chains**, without undermining scientific validity or violating pseudonymisation safeguards. Specifically, the Guidelines should:

- Distinguish between **rights exercisable at the point of care** (source records) and rights exercisable **at the level of the research sponsor**.
- Explain how rights apply when the sponsor **cannot re-identify** the data subject and therefore cannot directly fulfil access, rectification, or erasure requests.
- Clarify how Article 17(3)(d) and Article 21(6) apply in practice, including documentation expectations.
- Provide guidance for **registries, repositories, and secondary-use infrastructures**, where data subjects may not have direct contact with the controller.
- Encourage harmonised communication models to avoid contradictory expectations from ethics committees, DPAs, and research institutions.

### Concrete proposals

1. **Clarify that rights may be restricted under Art. 89 when necessary for scientific validity.**
2. **Provide examples of how to operationalise rights** when recontact is impossible (e.g., public notices, registry-level transparency).
3. **Clarify expectations for variant (in omics) reinterpretation:**
  - when return of results is required;
  - when it is not feasible;
  - how to document decisions.
4. **Provide guidance for group-level risks** (familial implications, hereditary traits).
5. **Provide operational models for exercising rights in research chains.** The Guidelines should include examples illustrating how rights can be fulfilled **without re-identification**, such as:
  - **Primary care provider / recruiting site as the natural contact point**
    - When the site acts as processor for the sponsor, it can fulfil access, rectification, or erasure requests **in the source medical record**, without disclosing the key code to the sponsor.

- When the site is an independent controller for care, it fulfils rights **in its own capacity**, while informing the sponsor only of necessary downstream implications (e.g., exclusion from future linkages).
- **Sponsor (private or public) or public-health institute fulfilling rights indirectly**
  - The sponsor acknowledges the request, informs the data subject that it cannot re-identify them, and **coordinates with the site** to implement feasible actions.
  - This preserves pseudonymisation and avoids unlawful re-identification.
- **Registries and repositories**
  - Provide a model where the registry acts as the **central rights-handling point**, even when data originate from multiple controllers.

**6. Clarify the application of Article 17(3)(d) and Article 21(6).** The Guidelines should explicitly state that:

- **Erase and objection may be restricted** when fulfilling them would “render impossible or seriously impair” the research objectives, provided Article 89(1) safeguards are in place.
- Controllers should **document this assessment** in the DPIA and in internal governance records.
- Controllers should offer **alternative mitigations**, such as:
  - exclusion from future linkages,
  - restriction of further use,
  - cessation of re-contact,
  - removal from future follow-up datasets.

Once data are pseudonymised and integrated into analyses, **individual deletion may no longer be technically feasible** without compromising scientific validity.

**7. Provide template language for transparency notices and Patient Information Sheet/Informed Consent Forms.**

The Guidelines should recommend standard wording for research contexts, including:

- **On pseudonymisation and limits of erasure**
  - *“Once your data are pseudonymised and integrated into aggregated analyses, it may no longer be possible to delete them individually without compromising the scientific validity of the research.”*
- **On who handles rights requests**
  - *“Your primary contact for exercising your rights is your treating physician or the site where you were enrolled. The research sponsor does not hold identifying information and cannot directly access your identity.”*
- **On remaining rights**
  - Clarify that access to general study results, rectification of source records, and complaint routes remain available.

**8. Clarify rights in secondary-use and health research.** The Guidelines should explicitly address:

- How rights apply when data are reused under Article 5(1)(b) compatibility presumption.
- How rights apply in **omics research**, where individual erasure may undermine population-level omics research and (re)use.
- How rights apply in **secure processing environments** (EHDS-data sharing route), where controllers may not have direct access to identifiers.



**9. Encourage harmonised national approaches.** The Guidelines should:

- Encourage DPAs to adopt **consistent interpretations** for rights in research.
- Recommend that ethics committees **do not impose contradictory requirements** (e.g., demanding erasure where pseudonymisation prevents it).
- Promote alignment with the other sector-specific regulations such as CTR, IVDR/MDR, EHDS, and the cross-sector DGA governance models.

## CONCLUSION

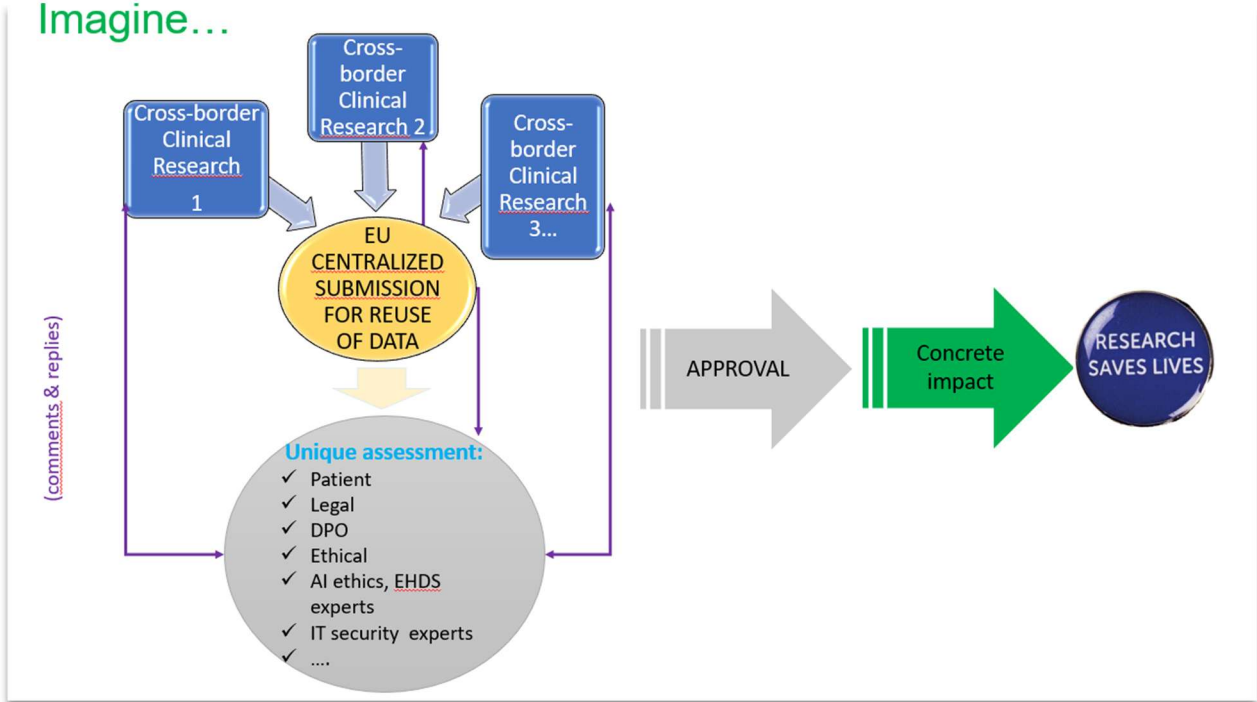
We see a strong need for a more coherent, centralised and future-proof approach to data governance for the reuse of health data in scientific research. While the Guidelines represent meaningful progress, a key element still missing is the practical operationalisation of the GDPR as it applies to health and scientific research. At present, the discussion remains centred primarily on privacy and data-protection concerns, whereas the broader challenge is to build a framework that enables research, validation, cross-border data reuse and algorithmic improvement in a lawful and proportionate manner.

Achieving this requires a shift in perspective. Research should not be treated as an exception to be constrained, but as a core public-interest activity that depends on clear, workable mechanisms rather than abstract principles. It also requires closer coordination across oversight bodies. The EDPB, EMA, national competent authorities, ethics committees, AI-ethics experts and health-data governance bodies all approach similar questions from different angles. Without shared principles, standards and operational definitions, fragmentation will persist.

Europe would benefit from the equivalent of Good Data Protection Practices for Health Research: a harmonised and jointly endorsed framework that clarifies roles, safeguards and lawful pathways for data reuse. Despite early concerns, the GDPR has not stopped or disrupted research in Europe. What persists instead is uncertainty, divergent interpretations and a lack of harmonised guidance, which creates significant operational burden for research organisations. Key concepts relevant to scientific and health research remain insufficiently defined, generating administrative work across privacy, regulatory and contractual teams without demonstrable benefit for data-subject protection.

To safeguard Europe's capacity for high-quality, ethically conducted health research, EU institutions, including the European Commission, the EDPB, DPAs, the EMA and national competent authorities, would greatly support the research community by providing coherent, harmonised and practically applicable guidance that enables lawful data use while ensuring proportionate and effective safeguards.

Imagine...



The EDPB Guidelines 1/2026 constitute an important step toward clarifying the application of the GDPR to scientific research. Yet, as demonstrated throughout this contribution, several structural gaps remain that, if left unaddressed, will continue to hinder lawful, high-quality, cross-border research across the Union. Research institutions, hospitals, clinical research sponsors, biobanks, registries, and emerging EHDS infrastructures operate in environments where research is longitudinal, multi-actor, omics-intensive, and dependent on the reuse of data collected over decades. In these ecosystems, legal certainty is not an abstract aspiration but a prerequisite for fulfilling statutory mandates, ensuring scientific integrity, and protecting the rights and interests of individuals.

Across the issues identified ( such as purpose definition, feasibility of consent, proportionality of safeguards, pseudonymisation, genetics and omics, international transfers, data-subject rights and the interplay with parallel EU legal instruments) a single cross-cutting message becomes clear: **the research ecosystem urgently needs agile, operational-focused, adaptable and fit to the research needs, and implementable legal clarity that reflects how scientific research actually functions in practice.**

The GDPR already contains the mechanisms to enable responsible research: Articles 6(1)(e), 9(2)(i) and (j), the presumption of compatibility, and the Article 89(1) safeguards. Yet the absence of consistent interpretation, combined with divergent national rules and ethics-committee practices, has produced a landscape where controllers face uncertainty, fragmentation, and procedural barriers that do not enhance data protection but do impede scientific progress. This is particularly acute in omics-intensive and rare-disease research, where identifiability risks are high, anonymisation is rarely feasible, and long-term, multi-actor infrastructures depend on the reuse of data collected over decades.

To address these structural gaps, the Guidelines should provide **clear, operational guidance** that supports harmonised, proportionate, and scientifically coherent implementation across Member States. This includes:

1. **Address persistent fragmentation across Member States** by clarifying divergent national laws, DPA interpretations, and ethics-committee requirements, particularly regarding secondary use, pseudonymization, legal bases, and data-subject rights in multi-actor research chains.
2. **Operationalise the GDPR for health and scientific research** by providing practical, operational guidance aligned with ICH-GCP, CTR, IVDR/MDR, EHDS and other established standards, and by developing Good Data Protection Practices or a Code of Conduct for Health Research.
3. **Clarify the interplay between the GDPR and other EU legal instruments** (CTR, IVDR/MDR, EHDS, DGA, AI Act) to prevent overlapping or contradictory obligations in public-health and clinical-research ecosystems.
4. **Provide a harmonised, research-specific interpretation of “purpose”** so that scientific research within a field or infrastructure is treated as one overarching purpose, with clear controllership roles, transitions under EHDS, and context-dependent safeguards for omics and longitudinal data.
5. **Acknowledge that GDPR consent-based models are not feasible** for long-term reuse, omics infrastructures, or linked datasets, and confirm Articles 6(1)(e) and 9(2)(i)/(j) with Article 89 safeguards as appropriate alternatives, while offering forward-looking guidance for consent under the EHDS.
6. **Clarify the distinction between GDPR consent and ethical consent** by avoiding dual tick-box systems and allowing a single integrated consent statement that reflects operational research reality and avoids confusing participants.
7. **Recognise the interdependence of research and research infrastructures** by treating registries, biobanks, secure platforms, and federated environments as integral components of scientific research whose long-term operation falls under the overarching research purpose.
8. **Clarify the cumulative application of Articles 6 and 9 GDPR** to ensure that combinations of legal bases do not circumvent EHDS safeguards for AI training, testing, and evaluation.
9. **Resolve unresolved complexities in research roles and responsibilities** by addressing investigators “wearing different hats,” varying sponsor–investigator models, and public-health institutes acting simultaneously as data holders, users, and authorities.
10. **Provide operational guidance on data-subject rights in research** to clarify how rights apply when data are pseudonymized, when multiple controllers are involved, or when re-identification is not possible.
11. **Adopt a risk-graded approach to genetic and omics data** instead of treating all genetic data as a single category, recognising the wide risk spectrum between WGS/WES, tumour-only somatic profiles, and single biomarkers.
12. **Provide research-tailored guidance on international data transfers** by establishing research-appropriate transfer routes for pseudonymised data when SCCs, TIAs, or adequacy

decisions are unavailable or impractical, including solutions for public bodies unable to sign SCCs, secure cloud processing, and federated data-visiting models.

13. **Clarify the cumulative legal-basis requirement under Articles 6 and 9 GDPR** in light of the EHDS, including how these articles apply to AI-related scientific research and how to avoid combinations that undermine EHDS safeguards.
14. **Recommend mandatory GDPR training modules** for ethics committees and Data Protection Authorities, including training on GDPR principles, clinical research, and sector-specific frameworks.

Strengthening the Guidelines along these lines would significantly enhance legal certainty, promote harmonisation across Member States, and ensure that Europe’s research ecosystem, including public health institutes, academic centres, clinical research sponsors, and EHDS infrastructures, can operate effectively, ethically, and in full compliance with the GDPR. A clearer, more operational, and more harmonised framework is not only essential for regulatory compliance; it is fundamental to Europe’s ability to generate the scientific evidence needed to advance personalised medicine, accelerate rare-disease research, and protect the health of its population.

**Contact details:**

**NoE on Omics Technologies: JANEWP9** [JANEWP9@sciensano.be](mailto:JANEWP9@sciensano.be)

- H el ene Antoine-Poirel [Helene.Antoine-Poirel@sciensano.be](mailto:Helene.Antoine-Poirel@sciensano.be)
- Inge Smeers [Inge.Smeers@sciensano.be](mailto:Inge.Smeers@sciensano.be)

**Task 9.3 - Legal and Ethical Aspects:**

- Wannes Van Hoof [Wannes.VanHoof@sciensano.be](mailto:Wannes.VanHoof@sciensano.be)
- Roxana Albu [Roxana.Albu@sciensano.be](mailto:Roxana.Albu@sciensano.be)