**European Network of Expertise (NoE) on Omics**

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

**Contribution to TEHDAS2 Public Consultation on**

**"Draft Guideline for Data Holders on making personal and non-personal electronic health data available for reuse"**

**(***Public consultation, November 2025***)**

Public consultations - Tehdas

## 1. PURPOSE OF THE GUIDELINE AND ITS RELEVANCE FOR THE OMICS COMMUNITY

The TEHDAS2 draft guideline (M6.1) provides practical and operational instructions for health data holders under the European Health Data Space (EHDS) Regulation. It clarifies the legal obligations of data holders, including timely provision of personal data upon permit approval, publication of anonymized or non-personal datasets, and submission of metadata to national catalogues. It also sets out workflows for data preparation, minimization, pseudonymization, anonymization, secure transfer, and post-delivery stewardship.

For the omics community, particularly in cancer genomics and multi-omics integration, this guideline offers a valuable baseline. It establishes a harmonized framework for secondary use of health data, provides clarity on roles and responsibilities of Health Data Access Bodies (HDABs), Trusted Data Holders, and Secure Processing Environments (SPEs), and introduces templates and checklists that can be adapted to omics datasets. The emphasis on auditability, transparency, and citizen-centric obligations is an important step towards building trust and ensuring compliance across Member States.

However, while the guideline is strong in defining generic EHDS compliance, it does not yet provide sufficient domain-specific guidance for omics-driven cancer research and clinical applications.

## 2. DESPITE ITS STRENGTHS, SEVERAL CRITICAL GAPS REMAIN FOR OMICS DATA HOLDERS AND USERS:

### Reuse of Data – Regulatory Interplay and Responsibilities

A central challenge in the European Health Data Space is determining whether data that were originally collected under other governing frameworks, such as the Clinical Trials Regulation (CTR) or the *In Vitro* Diagnostic Regulation (IVDR), can be lawfully reused for secondary purposes. The complexity arises because such data are subject to layered obligations: trial protocols, diagnostic validation requirements, GDPR safeguards, and now EHDS reuse provisions.

To ensure lawful and consistent reuse, clear guiding principles are needed to interlink these regulations. The responsibility for assessing whether reuse is permissible should not fall on a single actor but must be shared across the governance chain:

- **Health Data Access Bodies (HDABs)**: responsible for verifying that requests for reuse comply with EHDS obligations and do not conflict with other applicable regulations, such as CTR/IVDR restrictions.
- **Data Holders (labs, biobanks, hospitals, registries)**: responsible for documenting the original legal basis of data collection and flagging any reuse limitations arising from CTR/IVDR or other contractual clauses with the stakeholders involved in the initial data collection (trial conduct).
- **Data Controllers (sponsors, institutions, competent authorities)**: responsible for clarifying whether reuse aligns with the scope of the original consent, ethics approvals, and regulatory authorizations.

**KEY QUESTIONS TO ADDRESS AND MORE IMPORTANTLY: WHO SHOULD ADDRESS THEM?**

1. **Scope of reuse**: Is the data originally collected under other applicable regulations (CTR/IVDR) permitted to be reused under EHDS, and for which categories of secondary use? This question is far from straightforward. In practice, it is often impossible to state with certainty what reuse will aim at, because scientific inquiry is hypothesis-driven and evolves over time. Researchers may begin with one line of investigation but discover new directions as evidence accumulates. This inherent uncertainty makes rigid definitions of "secondary use" difficult to apply in a way that supports innovation while safeguarding trust.
2. **Consent and ethics:** Does the original consent or ethics approval cover reuse, or is additional approval required? At the same time, the issue of consent and ethics remains a constant debate. Does the original consent or ethics approval adequately cover reuse, or should additional approval be required when data are repurposed for new scientific questions? Because science is hypothesis-driven and evolves, it is often impossible to predict in advance the exact purpose of reuse. This makes proportionality, necessity, and actual relevance critical criteria for any balanced assessment. Ethics, authorities and supervisory bodies should provide clear guidance and robust recommendations that adapt to technological advances while ensuring continuity in research practices. Such guidance must help researchers understand when reuse is proportionate to the aims pursued, when it is necessary for advancing knowledge, and when it remains relevant to the original ethical commitments. At the same time, protective measures must be feasible, funded, and technically supported, so that safeguards are not only theoretical but operational. This balanced approach is essential to enable research that benefits patients and citizens, while ensuring transparency, continuity, and a harmonised interpretation and application of the law across jurisdictions.

    1. **Proportionality**: Reuse should be proportionate to the scientific aims pursued. The potential benefits for patients and citizens must justify the level of privacy risk and administrative burden involved.

    2. **Necessity**: Data reuse should be necessary for advancing knowledge or answering research questions that cannot be addressed otherwise. Authorities

should encourage researchers to demonstrate why the reuse is essential rather than optional.

3. **Actual relevance, ethics assessment, and mutual recognition:** Data reuse must remain relevant to the original ethical commitments and scientific purpose. Supervisory bodies should actively support researchers in assessing whether new uses align with the spirit of the original consent or ethics approval, or whether they fall under exemptions provided by the Helsinki Declaration or national laws governing data reuse. To ensure trust and continuity, such assessments must be harmonised. Decisions of ethics committees and the application of exemptions should be mutually recognized within the same country and across Member States. This approach avoids fragmentation, strengthens legal certainty, and enables researchers to pursue evolving scientific questions while respecting ethical boundaries and patient rights.

4. **Continuity**: Guidance should ensure continuity in research practices, so that evolving hypotheses and scientific directions can be pursued without constant procedural disruption. This requires flexible frameworks that acknowledge the dynamic nature of science.

5. **Harmonisation:** Interpretation and application of the law should be harmonised across jurisdictions. Regulators should provide clear recommendations that prevent fragmentation, ensuring that researchers can collaborate internationally without facing conflicting requirements.

3. **Regulatory overlap:** Can existing regulations like CTR and IVDR obligations coexist with EHDS reuse rules, or do they pull in different directions? The interaction between clinical trial and device regulations (CTR, IVDR) and the European Health Data Space (EHDS) reuse provisions exposes a critical tension at the heart of EU health governance. While CTR and IVDR prioritize participant protection, transparency, and regulatory oversight during data collection, EHDS introduces a paradigm shift, enabling broad secondary use of health data for research and innovation. This raises practical and ethical questions: Can consent given under CTR or IVDR be stretched to cover EHDS reuse? Who decides what constitutes “compatible” use? And how do ethics committees navigate this evolving landscape? These friction points, consent scope, reuse legitimacy, and ethics review, demand nuanced interpretation, not just legal precision. We present these dilemmas in plain language to spark dialogue among legal experts and stakeholders. When considered as a whole and meaningfully connected, these elements can shape governance that earns trust, reinforces accountability, and reflects proportionality.
4. **Governance responsibility:** Which actor (HDAB, data holder, data controller) has the authority to approve or deny reuse in complex, multi-stakeholder settings?
5. **Patient communication:** Who is responsible for informing the patient if incidental or clinically actionable findings arise during reuse, the data holder, the HDAB, the original controller, or the treating clinician?
6. **Cross-border harmonisation:** How can Member States ensure consistent interpretation of other applicable laws, such as CTR and IVDR, when determining reuse obligations under the EHDS, and what mechanisms can prevent fragmentation across jurisdictions?

**Standards and interoperability**: There is currently no reference to established FAIR omics standards, such as GA4GH APIs (htsget, refget, Beacon v2, Phenopackets), OMOP Genomics, ELIXIR-recommended metadata, or EGA/ENA submission standards. In addition, new standards should be developed and incorporated to address emerging omics domains, for example, OMPO-proteomics and related initiatives.

**High-dimensional privacy:** Current anonymisation guidance struggles with omics data because re-identification risks do not stem from the dataset alone. For those risks to materialise, two conditions must coincide: (a) disclosure of an individual's identity and (b) re-establishment of the link between that identity and the omics data. What is missing today is a framework that anticipates these dual conditions through genomic risk assessment, controlled access environments (SPEs), and trusted pseudonymisation services designed specifically for omics (Patient privacy in AI-driven omics methods - ScienceDirect, 2024).

**Data quality and harmonisation:** The guideline currently lacks operational frameworks for omics quality control, such as FastQC/MultiQC metrics, variant annotation standards (ClinVar, CIViC), batch correction procedures, and reference genome versioning. In addition, harmonization standards are urgently needed for emerging omics domains, such as proteomics and metabolomics, which are already entering clinical trials and routine use in several academic hospitals.

**Federated analysis:** No rules are provided for federated genome-wide association studies, federated variant curation, or federated machine learning. These are essential to avoid raw data transfers while enabling large-scale analysis.

**AI/ML governance:** The draft does not address documentation, validation, reproducibility, or bias assessment for AI models built from (omics) datasets. Training dataset versioning and transparency requirements are missing.

**Return of results:** While Article 61(5) of the EHDS Regulation references the obligation to communicate significant findings, the current draft guideline does not provide an operational pathway for managing genetic incidental findings in cancer datasets. In particular, there is no clarity on the allocation of responsibilities for recontacting patients, ensuring appropriate clinical communication, and defining the role of clinicians in this process.

Although some of these aspects are addressed in other TEHDAS guidelines, the linkage and procedural pathway should also be explicitly detailed in the guidance for Health Data Holders and Data Users. This is essential because the handling of incidental findings is inherently connected to their activities, roles, and responsibilities within the chain of data use and reuse. Are such data becoming "new data" to be shared? Who takes responsibility for their data governance under EHDS reuse of data? Without such integration, there is a risk of fragmented practices and uncertainty about accountability, which could undermine both patient trust and the operational feasibility of secondary data use in cancer genomics.

**Biobank and clinical integration:** The current guideline does not address the governance of biospecimens themselves, which are not directly covered under the EHDS Regulation. However, once samples are analysed, the resulting electronic data (genomic, proteomic, or other omics outputs) fall within the scope of EHDS and are held by data holders such as laboratories or biobanks. These entities often act as service providers, processing samples under the instruction of a data controller (e.g. hospital, research institution, sponsor). To ensure coherence, the guideline should explicitly recognise the role of sample identifiers and tracking tools in linking biospecimens to electronic datasets, and provide rules for how biobank samples, pathology images, and omics data are integrated into the EHDS ecosystem. Alignment with existing infrastructures such as BBMRI-ERIC, CanServ, and ESFRI should be included to avoid fragmentation and ensure interoperability across Member States.

**Cross-border harmonization:** There is no advice or guiding principle for harmonizing biobank and clinical-(gen)omic datasets across Member States, nor alignment with Cancer Mission initiatives which are already developed or on their way to be mature through EHDS implementations phases. Such guiding principles would benefit both patients and researchers and clinicians.

## 3. RECOMMENDATIONS FOR STRENGTHENING THE GUIDELINE

To ensure the guideline is fit for purpose for both the general health data community and the omics domain, the following recommendations are proposed:

**General recommendations for all health data holders:**

- Add annexes with operational tools such as a Data Acceptance Test checklist, re-identification risk statement template, and cost invoice breakdown template.
- Require opt-out linkage audits to prove capability of excluding opted-out records.
- Append minimum contractual clauses for intermediaries, Trusted Third Parties, and processors covering liability, audit rights, breach notification, and deletion/return.
- Publish national practice snapshots (UK Genomic Medicine Service, France Health Data Hub, Netherlands PALGA, Finland Findata, Belgium CAN.HEAL MTAs/DTAs) as exemplars.
- Propose guiding principles for the **interlink of different regulations applicable to reused data**, or **coordination decision tree**. This tool should ensure that data collected under different applicable laws can be assessed consistently for reuse under EHDS. Such a framework would provide transparency, reduce fragmentation, and support lawful, auditable secondary use of health data across Member States.
- Ensure HDABs have access to legal, scientific, ethics, and data protection expertise, supported by best-practice libraries and harmonized templates.
- Propose the establishment of **synergies and structured collaboration frameworks** that can serve as the backbone for a mandatory cooperative model between Health Data Access Bodies (HDABs) and data holders. Such

frameworks should ensure joint validation of **federated analysis workflows** and **data acceptance tests**, providing a consistent, auditable, and interoperable approach across Member States.

This collaborative backbone should:

- **Define shared responsibilities** and checkpoints for HDABs and data holders in validating federated workflows.
- Include **standardised** acceptance test protocols to guarantee data quality, reproducibility, and compliance before transfer.
- Facilitate transparent communication channels and escalation procedures to resolve discrepancies.
- Be anchored in harmonised templates and technical checklists, ensuring comparability and reducing fragmentation.

- By embedding these synergies into the guideline, the EHDS can foster trust, operational efficiency, and legal certainty in the reuse of health data, particularly in complex multi-centre and federated analysis contexts.

- Alongside the establishment of synergies and structured collaboration frameworks between Health Data Access Bodies (HDABs) and data holders, the guideline should also introduce a requirement for **mandatory training programmes** for all stakeholders involved in the permit decision chain. This includes HDAB staff, health data holders (HDHs), data users, ethics committees, and other relevant actors.

  Such training should:

  - Ensure a **shared understanding of the EHDS Regulation** and its interplay with other frameworks (CTR, IVDR, GDPR).
  - Provide practical guidance on **federated analysis workflows** and **data acceptance tests**, so that all parties apply consistent criteria.
  - Cover **ethical and legal responsibilities**, including patient communication obligations when significant or incidental findings arise.
  - Strengthen **cross-disciplinary collaboration**, ensuring that technical, legal, and ethical perspectives are aligned in permit assessments.
  - Be updated regularly to reflect evolving standards, interoperability requirements, and best practices across Member States.

  - Embedding mandatory training into the guideline would create a common baseline of competence and accountability, reduce fragmentation in decision-making, and foster mutual recognition, and build trust in the secondary use of health data.

## 4. ROLE OF THE JANE2 ELSI NAVIGATOR (SPECIFIC TO THE WORK IN JANE2)

**Specific recommendations for the JANE2 Omics community:**

- Mandate omics-specific metadata fields including provenance, reference genome, proteomics analytical and data pipeline meta data standards, QC metrics and pipeline versions. The guideline does not mandate omics specific metadata such as sample provenance (tumour/normal), reference genome version, sequencing coverage, extraction protocols, or quality control metrics. Without these, validation and linkage across centres are slowed.
- Align with GA4GH, ed ELIXIR, OMOP-Genomics, and EGA/ENA standards to ensure interoperability.
- Introduce genomic re-identification risk assessments and controlled-access SPEs for omics datasets.
- Provide frameworks for omics QC and harmonisation, including variant annotation standards, quantification standards for transcriptomics, proteomics and metabolomics, and batch correction methods.
- Establish rules for federated omics analyses (GWAS, variant curation, federated ML).
- Require AI/ML governance tools such as model cards, bias assessments, and reproducibility standards.
- Define pathways for return of incidental findings, with triage templates and clinician involvement.
- Harmonise biobank and clinical integration with existing frameworks, such as BBMRI-ERIC and CanServ standards.

The JANE2 ELSI Navigator aims to provide guidance that complements but goes beyond the TEHDAS2 guideline.

- **Module 1 – Regulatory Interplay Navigator**: This module maps overlaps and frictions between GDPR, CTR, IVDR, EHDS, and the AI Act. It helps omics stakeholders identify the correct legal route depending on whether activities are in research, validation, or clinical implementation. It supports proportional governance by aligning oversight with technology readiness levels. This is feasible and directly applicable to omics projects, offering decision trees and flowcharts that TEHDAS2 cannot provide within its generic scope.
- **Module 2 – National Ethics & Legal map**: This module captures divergences in national ethics approvals, IVDR exemptions, and GDPR interpretations. It provides country-specific checklists and trackers to align omics projects across borders. This is feasible and highly relevant for biobank and genomic datasets, addressing fragmentation that TEHDAS2 cannot resolve at EU level.

**Where the ELSI Navigator could add value beyond TEHDAS2:**

- Mapping national ethical/legal differences (outside EHDS scope).
- Detailed omics-specific regulatory interplay (CTR -IVDR- AI Act).
- Proportionality-based governance tailored to high-risk omics + AI contexts.

- Contractual frameworks for cross-border omics collaborations (Module 4).

## CONCLUSION

The TEHDAS2 draft guideline is clear, practical, and consultation-ready. It provides a strong baseline for generic EHDS compliance, balancing legal interpretation with operational recommendations. However, without additions on intellectual property coordination, federated analysis validation, governance decision trees across overlapping regulatory frameworks, opt-out proofing, contractual clauses, dataset quality checks, cost transparency, incidental findings pathways, re-identification risk metrics, and mandatory training for all stakeholders involved in permit decisions (including HDABs, health data holders, data users, and ethics committees), adoption risks fragmentation and unpredictability.

For the omics community, particularly cancer genomics and multi-omics integration, significant domain-specific guidance is required. The JANE2 Omics team recommends strengthening the guideline with omics-specific metadata, standards, privacy safeguards, QC frameworks, federated analysis rules, AI governance, incidental findings pathways, and biobank integration.

The ELSI Navigator developed under JANE2 WP9.3 can complement TEHDAS2 by providing regulatory interplay maps and national divergence heatmaps (Modules 1 and 2), filling gaps that the guideline itself cannot cover. Together, these tools will ensure predictable, auditable practice across Member States and provide a workable framework for omics consortia under diverse regulatory regimes.

*On behalf of the NoE on Omics in Cancer,*

**Contact details:**

**NoE on Omics Technologies: JANEWP9** JANEWP9@sciensano.be

- Hélène Antoine-Poirel Helene.Antoine-Poirel@sciensano.be
- Inge Smeers Inge.Smeers@sciensano.be

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

- Wannes Van Hoof Wannes.VanHoof@sciensano.be
- Roxana Albu Roxana.Albu@sciensano.be