EDPB Guidelines 02/2026 on Anonymisation

Farmindustria comments on Public Consultation

General Considerations

Farmindustria, the Italian association representing pharmaceutical companies, welcomes the EDPB Guidelines 02/2026 on Anonymisation and strongly supports their objective of clarifying the notion of anonymous data and providing a practical framework to determine whether anonymisation has been successful.

First of all, we note that Guidelines do not specifically address clinical research. It is therefore suggested that the document be supplemented by considering the specific characteristics of the sector, the various stakeholders involved, and the data flows typically associated with clinical trials.

In doing so, it is essential to preserve companies' discretion in assessing the appropriate measures, taking into account the specific circumstances and a risk-based approach, while avoiding any interpretation of the requirements as mandatory or any implementation that would impose disproportionate technical, organizational or operational burdens.

Legal Analysis – Section 2

In paragraphs 32 et seq. of Section 2 of the draft Guidelines under consultation, the EDPB is seemingly departing from the principles established by the Court of Justice of the European Union (CJEU) in Case C-413/23 P (EDPS v SRB).

In particular, the Court clarified that, when assessing the means "reasonably likely to be used" by third parties to re-identify an individual, "a means of identifying the data subject cannot reasonably be used where identification of that person is prohibited by law" (see paragraph 82 of the above-mentioned judgment).

By contrast, the approach proposed by the EDPB — which requires the controller to also assess potential unlawful conduct by third parties — makes the anonymisation process disproportionately burdensome and practically unfeasible, since unlawful activities, such as cyberattacks, are by their very nature unpredictable and *contra legem*. Consequently, assuming that the controller must prevent any potential criminal conduct risks undermining the very purpose and effectiveness of anonymisation itself.

It is therefore recommended that the Guidelines be aligned with CJEU case law, confirming that legal prohibitions, supplemented by appropriate technical and contractual safeguards, constitute sufficient elements to reduce the risk of re-identification to legally acceptable levels.

Furthermore, while it is acknowledged that prohibitions established in contracts between private parties do not constitute statutory provisions, it is suggested that the EDPB reconsider the possibility of recognising contractual prohibitions on re-identification. This assessment should also consider the purposes for which the data will be used by the third-party recipient (e.g. scientific research purposes) and the fact that the likelihood of re-identification must be "insignificant", rather than zero (paragraph 22).

Paragraph 21 – The rule whereby data remain personal data if "it cannot be ruled out" that the recipient has means of identification effectively introduces a zero-risk standard, which appears inconsistent with paragraph 22 and the SRB judgment. It would be necessary to rephrase this provision by referring to the means "reasonably likely to be used" by the recipient, based on objective factors.

Paragraph 38 – Regarding the legal basis for anonymisation, the presumption that it coincides with the legal basis for the original processing applies only where the purpose remains the same. In pharmaceutical practice, anonymisation is often carried out to enable secondary uses of data (research, data sharing, transparency).

It would be useful to clarify that anonymisation, as a risk-reduction measure, is generally compatible with the original processing under Article 6(4) GDPR, and to ensure consistency with Guidelines 1/2026 on scientific research, including about the derogations provided for under Article 9(2) GDPR.

The Technical Analysis of Anonymisation - Section 3

The Guidelines do not specify any metrics or thresholds for acceptable risk levels. In the pharmaceutical sector, established references already exist, such as the EMA guidelines on the anonymisation of clinical reports, which could be explicitly referred to.

Glossary - Section 4

It is suggested that a definition of “synthetic data” be added. It would also be useful to include an entry clarifying the relationship with EDPB Guidelines 01/2025 on pseudonymisation, specifying under which circumstances data that are pseudonymized for one entity may be considered anonymous for another one.

Relevant Entities

It should be clarified that, in the context of clinical research, the data recipient may also be the Sponsor and, where applicable, third parties involved in the clinical trial (delegated activities).

It should also be clarified that regulatory and supervisory authorities (e.g. EMA, AIFA, inspection authorities), which have access to data by law and are subject to confidentiality obligations, do not, by virtue of such access alone, render data personal where they would otherwise be considered anonymous for other recipients.

Means Reasonably Likely to be Used

The draft Guidelines already contain an important element of the criterion that we propose to further develop and emphasize, with specific, although not exclusive, reference to the clinical research context. Indeed, paragraph 31 provides that the potential value of re-identified data to the re-identifying entity should be taken into account to the extent that it offsets the costs and effort involved. However, the same paragraph 31 urges caution when relying on a "lack of motivation" as a factor, while paragraphs 92–93 highlight the increasing accessibility and decreasing costs of re-identification techniques.

This criterion should not be framed as a subjective test based on a "lack of motivation", but rather as an objective assessment of incentives (for actual re-identification) and proportionality. The time, costs, effort and other resources required for identification should be assessed in relation to the reasonably foreseeable value or utility arising from the actual identification of the individual, together with the legal, technical and organisational barriers and the risks faced by the entity undertaking re-identification.

In the specific context of clinical research, of Real World Evidence and irreversibly anonymised datasets, for a pharmaceutical company sponsoring a clinical trial, the scientific and economic value of the dataset will normally lie in the analysis of clinical outcomes, safety, efficacy, treatment patterns and other variables at cohort or population level. By contrast, the actual identity of an individual patient will generally provide no relevant scientific or economic value for these purposes. Since the sponsor does not hold additional identifying information and has no role in the patient's treatment, identifying a specific patient would require a disproportionate and entirely irrational investigative effort, without producing any reasonably foreseeable benefit. In such circumstances, the mere

demonstration that re-identification is technically conceivable, potentially by combining data with external sources or through a "chain of means", should not make identification "reasonably likely".

In light of the above, it is proposed that paragraph 29(d), paragraph 31 and the contextual assessment set out in Section 3.5.2 of the Guidelines expressly require consideration of the relationship between (i) the costs, time, effort and practical barriers associated with identification and (ii) the concrete and objectively foreseeable benefit or utility of identification for the relevant entity. The same factor should also be reflected in Annex 1 ("Flowchart for the technical analysis of anonymity").

This proposal is consistent with the overall approach of the Guidelines and with Recital 26 GDPR.

It would give full effect to the distinction — already recognised in paragraphs 8 and 22 (according to which identifiability is a matter of probability, and that probability must be insignificant in practice, although not necessarily zero), 26 and 49 of the Guidelines — between theoretical possibility and a significant likelihood of identification in practice, while preserving the EDPB's legitimate caution that subjective intentions alone are neither sufficient nor stable and consistent over time.

In other words, we propose introducing an objective criterion into the assessment, which could be referred to as an "incentive and proportionality factor". This would not replace the other factors but would make the distinction between theoretical possibility and significant likelihood in practice, as required by Recital 26 GDPR, more effective.

Specifically, the EDPB is invited to draw the logical conclusions from its statement concerning the "potential value" in paragraph 31, according to which:

"When assessing the means reasonably likely to be used, the potential value of the re-identified data to the re-identifying entity (which is not limited to monetary value) should be taken into account to the extent that it offsets the costs and effort involved in re-identification".

If the existence of a benefit that offsets the costs makes the use of means of re-identification more plausible and likely, the objective absence of any benefit, in the face of non-negligible costs and efforts, must be capable of weighing in the opposite direction.

Based on experience in the field of clinical research, coded data received by the sponsor generally continues to be considered pseudonymized, even where the re-identification key remains with the clinical trial site.

However, specific clarification is needed regarding the various actors involved — sponsors, CROs, central laboratories, technology providers and independent researchers — and their actual ability to cooperate with the entity holding the key.

A case-by-case assessment of the actual possibility of identifying an individual is therefore appropriate.

For example, consideration should be given to whether the data recipient can access additional information or the necessary re-identification code, as well as the time, costs, expertise required, legal and contractual restrictions, and existing security measures.

Contextual Approach vs Simplified Approach

It is suggested that the Guidelines clarify that the simplified approach should only serve as an initial assessment.

The theoretical existence of a technical method for identifying an individual does not automatically mean that the data cannot be considered anonymous.

It is necessary to assess whether, in specific circumstances, that method is genuinely likely to be used.

The "contextual approach" proposed by the EDPB is welcomed, as it distinguishes theoretical possibility from the means realistically likely to be used, considering both evolving technological capabilities and decreasing technical costs. We welcome the recognition, in paragraphs 26–31, that the assessment of the "means reasonably likely to be used" to re-identify an individual requires an objective evaluation based on the actual likelihood of re-identification.

Of the two approaches, the contextual approach is considered preferable, as it makes it possible to distinguish between:

- the clinical trial site, which holds the subject-patient identification code and can link the data back to the data subject;
- the sponsor, which normally receives coded (pseudonymized) data;
- any third-party recipients, which may have an even more limited level of access.

No Record Isolation

The Guidelines should clarify that a unique data record or an isolated case does not automatically make an individual identifiable. In oncology trials, in CAR-T therapies, rare diseases, genomic studies and small patient populations, a patient may be statistically unique without it being practically possible to link the record to their identity.

This is a particularly important point for the pharmaceutical sector.

In clinical datasets, almost all records are unique, meaning that the criterion would almost always be breached. Example 24 clarifies that uniqueness is not sufficient to rule out anonymity. It would be useful to include a similar example in the healthcare context, distinguishing between attributes that can be matched with external information (age, sex, geographical area, dates) and clinical attributes that are unlikely to be available from external sources.

No Inference

With regard to scientific research, the EDPB should distinguish between general conclusions concerning a group of individuals and conclusions relating to a specific individual. The latter should be relevant for the purposes of anonymisation only where they allow information to be linked with sufficient certainty to an identified or identifiable individual and may have concrete consequences for that individual.

Periodic Reassessment

With regard to paragraph 35 of Section 2, while the principle that technological developments may lead to the emergence of tools capable of overcoming the security measures adopted for data anonymisation is acknowledged, it should be considered that, once data have been provided in anonymous form, they fall outside the scope of the GDPR and are fully under the control of the third party receiving them. Consequently, any periodic reassessments may prove ineffective in practice.

Paragraph 35 implies that anonymous data may subsequently "become" personal data again, potentially giving rise to liability for the controller. However, in the case of data that have already been published or shared (e.g. clinical reports or datasets made available to the scientific community), the controller can no longer intervene.

It would therefore be useful to clarify that the controller's responsibility should be assessed in light of the due diligence exercised at the time of disclosure, and that periodic reassessment should concern only data that remain under the controller's control.

Anonymisation as a Process

We agree that anonymisation constitutes a processing activity subject to the GDPR (paragraph 38). It would be useful to clarify whether and under what circumstances a Data Protection Impact

Assessment (DPIA) is required for the anonymisation process, as well as the role of a service provider carrying out anonymisation on the controller's instructions (as a processor under Article 28 GDPR).

Rare Diseases / Small Patient Populations

The Guidelines should specifically address cases involving rare diseases and small patient populations. In these contexts, certain data may be unique precisely because of the limited number of patients involved. This factor alone should not prevent data from being considered anonymous where the recipient does not have practical means of linking them to a specific individual.

Based on the approach described in Section 3 of the Guidelines, where data intended to be considered anonymous relate to small populations or individuals affected by rare diseases, the strict application of the technical approaches proposed by the EDPB (whether contextual or simplified) would almost inevitably lead to the data being classified as "personal data".

In this regard, it would be useful to broaden the assessment of the effectiveness of anonymisation by also taking into account the specific purpose of the processing (e.g. scientific research), together with the relevant organizational and contractual safeguards (e.g. strict prohibitions on re-identification, segregated research environments and access restricted to authorized personnel only).

Special Categories of Data (Genetic and Genomic Data)

- Genetic and genomic data constitute a category of data presenting a high risk of re-identification due to their uniqueness, permanence and potential to be linked to other sources of information. The effectiveness of anonymisation must be assessed on a case-by-case basis and periodically reassessed in light of technological developments and the availability of new data sources.
- Based on the approach described in Section 3 of the Guidelines, where data intended to be considered anonymous relate to small populations or individuals affected by rare diseases, the strict application of the technical approaches proposed by the EDPB (whether contextual or simplified) would almost inevitably lead to the data being classified as "personal data".

  In this regard, it would be useful to broaden the assessment of the effectiveness of anonymisation by also taking into account the specific purpose of the processing (e.g. scientific research), together with the relevant organizational and contractual safeguards (e.g. strict prohibitions on re-identification, segregated research environments and access restricted to authorised personnel only).

Other Points

Although the data transferred to the Sponsor are classified as "pseudonymized" in absolute terms, since the re-identification key is held by the clinical trial site, they may be considered anonymous from the Sponsor's perspective where the latter neither possesses nor can reasonably access the means necessary to identify the data subjects.

In this context, the ability to re-identify individuals rests exclusively with the clinical trial site, which maintains the subject identification list and remains the only data controller capable of linking the code to the patient's identity.

This assessment is consistent with the contextual approach advocated by the EDPB in Guidelines 02/2026 and with the developments in case law that inspired them, according to which anonymity must be assessed from the perspective of the relevant entity receiving the data.

In light of Example 2 (SRB), it would be useful to clarify whether Chapter V of the GDPR applies where the data are anonymous from the perspective of the recipient in a third country (e.g. a parent company) but remain personal data for the exporter.