Response to Public consultation on EDPB **Guidelines 1/2026 on processing of personal data for scientific research purposes**

Dr. Maria K. Andersen Date: 24.06.2026	**Declaration:** This feedback is provided solely in my personal capacity, as an individual and experienced researcher. It should not be interpreted as representing the official views or position of the University that I am employed by.

As the EDPB must already be aware of, the implementation and practice of GDPR within research has been highly fragmented and inconsistent across Europe. These inconsistencies go beyond national laws, as even similar institutions (e.g. Universities) within the same country often apply differing interpretations of GDPR in research. **The work of EDPB to provide dedicated guidelines for scientific research is therefore highly welcomed and appreciated.**

In this feedback I will start with describing **my own experience as a researcher** working with clinical and biomolecular data. This will provide context to the real-life challenges researchers face concerning GDPR. I will then move on to my **suggestions for improvements** as summarized here:

1. Clearer guidance on balancing GDPR with **Open Science** principles
    a. Open archiving of research data
    b. Information to data subjects to enable future open science compliance
2. **Overly restrictive language** in the guidelines defeats consistent practices
3. Addressing challenges of using Consent as the legal basis (Article 6(1)(a))
    a. Need more clarity on whether the rights of the data subjects will differ based on the legal basis used pursuant on Article 6(1)
    b. Get rid of '4.1.2.2 Dynamic consent'
4. More clarifications on when **erasure** will seriously hinder the scientific purpose
5. Suggestion on how to practically apply **right to object – "Opt-out Consent"**
6. Guidance on how the **controller** can establish that it is the organization, not the individual researchers, which are legally responsible for the duties of GDPR
7. More reflection on the **risks for scientific integrity** by uncritically promoting anonymization, deletion and data minimization

0. My personal experience with GDPR as a biomedical researcher

My educational background is molecular biology, and I have been working with so-called 'sensitive data' from prostate cancer patients for 12 years. Today, I am the project leader, or principal investigator (PI), of two funded projects within prostate cancer research. Within our research environment here at the Norwegian University of Science and Technology (NTNU), we work with a wide range of data types, including clinical data manually collected from patients' journals, clinical MR images, molecular omics data

(RNA-sequencing, proteins, metabolites, DNA methylation), histology images and data from Norwegian national health registries (prescription drugs, cancer registry).

There is a broad consensus among myself, my colleagues at NTNU and research collaborators both within and outside Norway that our research data must require careful handling. Ensuring the confidentiality and trust of our research participants is of high importance. I (and my colleagues) spend considerable extra time and resources because we do research on sensitive personal data. Much of this extra work is both reasonable and necessary to protect the rights and privacy of our research participants. I will not spend time describing the activities I do that are reasonably part of my job. Instead, I would like to address the time and energy wasted (of myself and others) due to unreasonable interpretations of GDPR and having to deal with conflicting requirements of Open Science principles:

Experience A) I have spent several workdays arguing with editors (of scientific journals[1]) on why we cannot share any research data openly, despite that I personally think it would be possible to share some of the processed data without risking re-identification of participants. All because we still have the link to the patients in the form of a re-identification key[2] (subjected to high-level safeguards, not to be shared).

Experience B) I have in total spent several work weeks communicating with other researchers requesting access to our data. Explaining to them the lengthy process of getting access to this data, determining purpose compatibility based on what participants were informed about, that we would need ethical approval to share, DPIA needs to be updated and approved by several joint controllers, there will be back-and-forth on the data sharing agreement and they may need organize remote access to the data. All of this regulatory workload despite that I know that it would be possible to share minimized and processed versions of the data that for all intents and purposes is anonymous for them, and for the rest of the world.[3]

Experience C) I know of several cases where researchers have spent months requesting data for research from national health registries, spending unreasonable efforts to argue and justify to administrators (who are prioritizing data minimization above all else) why they need the variables they are asking for.[4]

---

[1] Many scientific journals now require that the underlying research data is made available. Saying "data available upon request" is not accepted any more.

[2] Various Norwegian authorities and researching institutions have largely and uncritically adopted the definition of anonymization introduced by Working Group 29 in Opinion 05/2014 on Anonymization Techniques, which conveys that anonymization must be irreversible, even for the controller.

[3] In cases where we have ended up sharing data, it has become a lengthy process of drafting data sharing agreements, back-and-forth between lawyers at the respective universities, figuring out and setting up technical solutions for sharing the data.

[4] It can be challenging to explain to someone without any experience in research why a seemingly unrelated piece of information (e.g. annual income) can still have a confounding effect when analyzing clinical treatment response and may need to be controlled for in the statistical analysis.

This is despite the documented insurance of extra levels of safe-guards (external ethical approval, pseudomyzation, high level secure storage, and access control to a few researchers), and that participant consent covers using the registry data.

Experience D) In an attempt to educate myself on how to balance the requirements of Open Science with GDPR-compliance, I have over the years sought out numerous training sessions, written materials, webinars and courses on Open Science and data management for research data with personal data. To my frustration the most concrete advice I'm always left with is "as open as possible, as closed as necessary". A statement that tells me absolutely nothing.

Experience E) I know of several cases where researchers were told they could not anonymize[5] their dataset because they unfortunately had not mentioned the possibility of anonymization in the original consent form. Supposedly, the "anonymization process is still processing of personal data".[6] They were instructed that they would need to obtain a new active consent from every participant (an impossible task as I will get back to).

A general remark: It is clear from reading the Guidelines as a whole that the authors are far more familiar with how privacy and data protection are dealt with in a) health research and b) very large population studies. This is understandable given the work from EDPB on the EHDS, the fact that health and clinical research is more often regulated by laws and regulations (the mother tongue of lawyers[7]), and that most information on suggested models on how to handle privacy/GDPR in research, including consent, routinely updated research information and systematic ways to handle request for data portability or erasure, comes from comprehensive and large study population cohorts. However, the EDPB and its Secretariat should keep in mind that most research projects in Europe are much smaller, like mine. We do not have the resources to follow the suggestions in these guidelines on sending out regular newsletters to each participant, developing a privacy dashboard or implementing "dynamic consent". When revising these guidelines, the EDPB would be wise to collaborate more closely with researchers managing smaller (average-sized) projects and researchers from non-medical fields.

[5] Here anonymize according to the old definition from WG29 requiring deletion of re-identification key

[6] I suspect this is a wrongful application of the statement in WG29 Opinion 1/2014 "Anonymisation constitutes a further processing of personal data".

[7] From what I can gather, EDPB and its Secretariat is dominated by members and experts with a legal background.

1. The battle between GDPR and Open Science principles

The European Union has increasingly been prioritizing **Open Science** as a core pillar of its research strategy to accelerate knowledge sharing and transparency. Open science is even a legal obligation for projects funded by Horizon Europe. A substantial part of these Open science principles concerns open data, including open sharing of data, reuse of data and enabling verification of research results. Given the strong focus on open science in EU and globally, I find it highly surprising, and somewhat worrisome, that open science was not mentioned once in the guidelines.

There is an urgent need for practical and concrete advice that goes beyond "As open as possible, as closed as necessary", a phrase that tells me *absolutely nothing* about what data and in what formats would be acceptable to share openly and/or with collaborators. The guidelines should have a whole section dedicated on how to best balance GDPR with the principles of open science, in particular in regards to sharing research data.

Openly archiving anonymized versions of research data while retaining identification key

Paragraph 162 in the guidelines appears to convey that research data containing personal data (or information derived from personal data) can exist in different formats, and can be considered anonymized to a receiver without access to additional information required to reasonably be able to re-identify the research participants. This is a notion I am very happy to finally see from the EDPB.

Based on this logic, it should be acceptable practice to openly share versions of the data that nobody (other than the controller who has the reidentification key[8]) would be reasonably able to identify the data subject from. The Swedish National Data service (SND) previously provided highly relevant questions and suggestions to the EDPB in this regard[9]. I suggest that the guidelines include a statement similar to:

> "If a controller creates a version of a dataset that has been anonymised such that recipients, including the general public, cannot reasonably re-identify data subjects, that version may be shared openly, even if the original controller retains a identification key to participants (the data subjects)."

Recommendations on elements to include in the privacy statement that enables adherence to Open Science principles

Given that Open Science principles and secondary re-use (further processing) of research data is increasingly wanted within the EU and scientific community, the EDPB should issue guidelines and examples of how to construct the information notice to

[8] On the criterion that the re-identifications key is subjected to strong safe-guards pursuant of Article 89
[9] Feedback reference 01/2025-0010 from SND on EDPB Guidelines 01/2025 on Pseudonymisation

research participants that enable such practice. It should include information that will cover:

- Potential future sharing with other researchers
- Archiving research data; open access and/or restricted access
- Re-use of data for new research projects not known at the time of collection
- Sharing data for the purpose of peer review

2. Overly restrictive language defeats the purpose of implementing consistent practices

I note that the EDPB in these guidelines, as well as other guidelines and opinions, often choose to use stricter language than GDPR and its recitals. For example "strictly necessary" is often used by EDBP, a formulation that does not appear in any of the GDPR articles. While I appreciate that this restrictive choice in language is rooted in EDPBs objective to reduce the risk of breaking the rights of data subjects, this can have a damaging effect on the quality and integrity of scientific research.

For example, within the theme of data minization;

a) 'Strictly necessary' may cause researchers to only collect data they know for certain is needed in the project, and exclude important variables they suspect could be relevant.
b) 'Necessary' opens up to collecting data that could be relevant for the research question.

Research is explorative in its very nature and it is often not possible to anticipate down to the exact details of what variables (and at what detail level) will be needed to secure trustworthy and high-quality results. Linguistically, "strictly necessary" introduces far more uncertainty and subjective interpretations than "necessary". Hence, resulting in more inconsistent practices on what is acceptable or not, and would go against the primary mandate of the EDPB which is to ensure consistent application of GDPR (Article 70(1)). The situation described in **Experience C** illustrates the risks of too restrictive language.

I suggest that the guideline remove all phrasings that suggest a more restrictive application then the GDPR and its recitals:

Para-graph	EDPB restrictive language	GDPR / recital phrasing	Suggestion
29	"Storage for generic scientific research purposes without any specification of purpose cannot be justified with reference to Article 5(1)(e) GDPR"; future research	Art. 5(1)(e) says data may be stored for longer periods where processed solely for scientific research purposes, subject to Art. 89(1) safeguards; it does not expressly require "reasonably foreseeable" future projects or prohibit	Rewrite paragraph 29.

	should be "reasonably foreseeable".	"generic" research storage in those words.	
96	Art. 14(5) exceptions "must be interpreted restrictively and should not be relied on routinely."	Art. 14(5) says information duties do not apply where one of the listed exceptions applies, including impossibility, disproportionate effort, serious impairment of research objectives. Within some research fields or for some data sources, these conditions are very often met and this derogation is used routinely.	Remove paragraph 96
119	Necessity in Art. 17(3) "must be interpreted restrictively"; controller must show processing is "strictly necessary" for an exception.	Art. 17(3) says the erasure right does not apply where processing is "necessary" for listed grounds.	Replace "strictly necessary" with "necessary"
120	Art. 17(3)(d) "should only be applied in limited circumstances"; "render impossible or seriously impair" implies a "high threshold".	Art. 17(3)(d) says erasure does not apply where processing is necessary for scientific research in accordance with Art. 89(1), insofar as erasure is likely to render impossible or seriously impair achievement of the research objectives.	Replace "strictly necessary" with "necessary". Remove statement on "limited circumstances".
126	Necessity under Art. 21(6) "must be interpreted restrictively"; processing must be "strictly necessary" and the task cannot be carried out, or not as efficiently, without the data.	Art. 21(6) says the data subject may object to research processing unless the processing is "necessary for the performance of a task carried out for reasons of public interest."	Replace "strictly necessary" with "necessary".
154	Controllers "must determine which data are strictly necessary to process."	Art. 5(1)(c) says data must be "adequate, relevant and limited to what is necessary".	Replace "strictly necessary" with "necessary"
159	Directly identifiable data should only be processed where "strictly necessary and proportionate" for the research purposes.	Art. 89(1) says safeguards may include pseudonymisation where purposes can be fulfilled that way, and where purposes can be fulfilled by processing that no longer permits identification, they shall be fulfilled that way.	Replace "strictly necessary" with "necessary"

3. The practical consequences and challenges of consent (Article 6(1)(a) GDPR)

Consent as a legal basis has been proven challenging within a research setting. The strict requirements of using GDPR-consent is often not compatible with the needs to research. Many are not aware of these strict requirements, thinking it is the same as the ethical consent to participate[10]. Some think consent is the preferred legal basis (even where

[10] Something these guidelines provide more clarity on in Section 4.1.3.

other bases legally can be chosen) because it is typically presented first in GDPR (Article 6 and 9) and in most guidelines. Even the fact that the EDPB guidelines dedicate 7 whole pages to consent, but less than one page to public interest and legitimate interest may signal that consent is the preferred legal basis.

However, for me, the lengthy discussion and examples from the EDPB on consent and all the different things one has to consider, demonstrates that using consent as the legal basis for research projects is just a dumpster-fire of complications and confusion. **The guidelines should take a clearer stance that when union or MS laws support it, public interest (Article 6(1)(e) GDPR) should be the preferred and more practical legal basis** (followed by legitimate interest). This will ensure a more consistent application of GDPR within a sector that perhaps face the most challenges in implementing GDPR in a consistent way.[11]

Clarifications of what a withdrawal of consent actually mean: To delete or not delete

The guidelines should clarify exactly what should happen if a research participant withdraws their GDPR-consent. I have heard conflicting interpretations of what should happen. Many privacy experts say that all collected data has to be deleted if consent was used as the lawful basis. However, from the GDPR articles 7 and 17, it is not obvious that deletion (unless other lawful basis) will be the default option in every scenario:

Article 7(3) states as a condition for consent "[1]*The data subject shall have the right to withdraw his or her consent at any time.* [2]*The withdrawal of consent shall not affect the lawfulness of processing based on consent before its withdrawal.*"

Article 17(1)(b) GDPR: "*the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing*;"

The conditions for consent set out in Article 7 together with Article 17(1)(b) points towards that all collected research data needs to be deleted if consent is withdrawn. At least this is how I have seen this interpreted by several legal experts across different entities in Norway and by the Norwegian DPA (Datatilsynet). There is also confusion on whether "legal ground for the processing" refers to using another legal basis pursuant of Article 6(1) or if the legal grounds can be a sperate union and/or MS law that mandates or allows data retention. **The guidelines should clarify what is meant by "other legal grounds" in this context.**

[11] As EDPB states in **paragraph 3** "the processing of personal data in the area of scientific research is diverse and complex, not least as there are many different actors, fields of research, types of data and categories of data subjects involved, as well as different technologies and methods used to collect and process data."

However, Article (17)(3)(d) provides an exception if the purpose is scientific research: "*or archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing;"*

The exception is further emphasized in Recital 65: "*However, the further retention of the personal data should be lawful where it is necessary, for... [...] scientific or historical research purposes or statistical purposes...*"

Interpreting GDPR Article 7 and 17 together makes it sound like, if the purpose is scientific research, it makes no difference if the legal basis is (a) consent, (e) public interest or (f) legitimate interest. It will be the same criteria and assessment that needs to be done; ie whether the erasure will seriously impair the achievement of the scientific project**. The Guidelines should clarify whether, once Article 17(3)(d) is invoked, the practical threshold for refusing erasure is the same regardless of whether the original Article 6 basis was consent, public interest, or legitimate interest.**

Dynamic consent should just be removed from the guidelines

First of all, the guidelines appears to give conflicting descriptions of what dynamic consent is. In the **Executive summary** and in **paragraph 41** and **52** it sounds like any individual sub-project should preferably aim to collect a new consent specific for that project, even if the project is covered by the original broad consent. In contrast, **paragraph 46** and **51** gives the impression that dynamic consent comes into play when the research project fall outside what was covered in the broad consent. To me, the last scenario just sounds like the general rule of needing to update consent when processing is no longer covered by the original consent, which applies regardless of the consent being broad or specific. So why call this dynamic consent and not just update of consent or separate consent for a new project?

If the guidelines however intended to convey the first scenario, it should be noted re-contacting data subjects to obtain new consent will often be highly impractical or infeasible in most research situations. The only scenario I can think of where dynamic consent would work in practice is, as rightfully described in **paragraph 52**, where "researchers are in a close and prolonged relationship with the data subjects" and "where researchers interact with research participants on a regular basis in long-term research projects". Research projects where there are regular interactions over time with participants are typically more specific in nature and therefore does not need to use broad consent in the first place.

It should be noted that 'Dynamic consent' as implemented in large real-life population studies, differs from how these guidelines describe dynamic consent. Dynamic consent is implemented as comprehensive technical infrastructures with much broader applications that cover granular consent options, easier management of revisable consent choices, dynamic communication of study progress and recruitment to collect more data through participation in surveys. It is something else than 'Consent as a legal basis' within GDPR. It is therefore inappropriate to have a section called '**4.1.2.2 Dynamic consent'** within '**4.1 Consen**t' alongside with section '**4.1.2.2 Broad consent**' (which is a real GDPR consent thing) as it gives the wrong impression. I therefore recommend that the section and recommendation of dynamic consent is removed in its entirety.

4. Severe consequences of erasure in research projects are not limited

Paragraph 120 states: "..Article 17(3)(d) GDPR should only be applied in limited circumstances, as it is only when erasure is likely to render impossible or seriously impair the achievement of the scientific research purposes that it is possible for a controller to reject a request for erasure"

"only be applied in limited circumstances" gives the impression that it would be highly unusual or unlikely that erasing part of the research data would seriously impair the research project. EDPB does provide very useful context further in paragraph 120 on how erasure may negatively impact the research quality to a much higher degree for studies with few participants as compared to larger studies. However, I must point out that there is important context of how scientific research based on datasets works that also makes it very relevant *when* a request for erasure is received. If the erasure is requested at a late enough time-point in the research project, it will in practice be likely to seriously impair the achievement of the purpose, i.e. the research project.

The concept of a dataset: In research, the personal information of individual participants together make up *a* dataset. The dataset is a package that as a unit produce one or more scientific results. If data of even one participant is removed, the exact result will change.

This means that for a research project where the data analysis is completed, the whole data analysis will need to be re-done if one is required to delete data from one participant. The quality of the research results may not be severely impacted if the study cohort is large, but the project as a whole can in many circumstances be delayed to such degree

that it will not be completed[12]. The table below shows how the timepoint of receiving the erasure request impacts a typical project

	Phase of research project	**Consequence of erasing data for achieving the scientific purpose and adhering to scientific guidelines**
1	Recruiting participants and collecting data	**Low** – A good study design will have taken into account potential withdrawal of participation. It may be easy to find replacements while still in the recruitment and data collection phase
2	Preprocessing and structuring data	**Medium/low** – In case of large datasets, deletion of 1 data subject will have limited impact
3	Analysis of data to produce results	**Medium** – Same as above (stage 2). Data analysis will need to be redone, which can cause some delay.
4	Interpretation of results and writing manuscript	**Medium/high** – All data analysis and documentation of results will need to be redone. Dissemination of results is delayed
5	Submit manuscript for publication	**High** – All data analysis and documentation of results will need to be redone.
6	Peer review	**Very high** - All data analysis and documentation of results will need to be redone. Reviewers may need to redo their assessments
7	Published results	**Maybe illegal** – Although one is not expected to redo the analysis at this point, deleting data will remove the chance to verify results. Archiving all data may either be mandated by law or if not good scientific conduct requires data retention.

The guidelines should also address time-point of the request in relation to the project progression and to what degree the erasure would delay the project.

Example 19 should be replaced: This is a strange example which is not even a request for erasure, but a request for rectification (or should be). And in that regard, it is also unclear to me why it would severely impact the research to change the name based on the request. Most research projects can rely on pseudonymization anyway. As long as all data entries are connected to the right individuals, they can be called whatever in the dataset. In this example, the data is also publicly available (by someone else). So why would the requestor need the researchers to change or delete his name in their research processing when his old name was publicly exposed in the public data anyway?

[12] Many research projects in the university sector are driven by individual researchers that have been given finite external funding which also covers their employment and/or the employment of other scientific staff.

5. Right to object-section is not very informative.

Section 6.3 on the right to object reads more legalistic compared to the other parts of the guidelines. This part is currently not very informative on how Article 21 should be applied in practice in a research setting. It would also be useful to add examples of how this right comes into play in a research project.

Suggestion: Bring back the passive/opt-out consent

Prior to the GDPR coming into effect, it was common in Norway (and likely other countries) to use so-called 'passive consent' or 'opt-out' consent. This solution was typically used in situations where the researchers were not in direct contact with the participants, but using their personal data in their research. The participants would receive information on the project and the planned use of their data and would be given an absolute guarantee that personal data will not be processed if an objection is given within a reasonable timeframe (for example 30 or 60 days). The researchers would wait to start collecting and processing the data until the given time-frame had passed. After that, erasure and further processing would depend on whether the erasure will severely impact the completion of the project (evaluated on the same basis as if a signed consent was given).

The opt-out consent was a very pragmatic and ethical solution that balanced the need for efficient project management and giving the participants proper information and the right to object and not be included. **I suggest the guidelines could re-introduce this opt-out recruitment strategy within the framework of Article 21[13].** In a research setting, it is possible, and even normal, to afford more rights to the data subject than Article 21 demands. This includes that the potential research participants are given a time period to object before any processing start and they would not be required to give any reason for their objection[14].

6. Need clarifications on who is the controller

Many research projects in the university sector are initiated and driven by individual researchers or PIs (principal investigators). From the perspective of the researchers, they are the ones who determine the purpose and the means of the processing (Article 4(7))[15]. In practice, they are also the ones who evaluate and determine the legal basis to be used

[13] For project that have used Article(1)(e) and (f) as the legal basis

[14] Article 21(1) GDPR states that "data subject shall have the right to object, on grounds relating to his or her particular situation". This implies that the data subject must provide a reason for the objection. As this is typically not necessary and sometimes viewed as unethical in research, research participants are informed specifically that they can ask for erasure (or object) without providing a reason.

[15] For ERC-funded research projects, the PI has even has the power to move the project, and hence processing of personal data, to any other university in EU. This further creates uncertainty on who is the controller

(Article 6(1), and Article 9(2) if applicable), organizes informing the research participants (Article 12-14), handles requests for erasure (Article 17), ensures proper safeguards are in place (Article 89(1)), performs the risk assessment (Article 32(2) and Recital 76), determines if a DPIA is required and filling out the DPIA (Article 35), write and organizes agreements if sharing data with research collaborators and are responsible for deleting the data when the research is concluded. All tasks a controller is responsible for according to GDPR. Universities do typically create guidelines on GDPR practices (Article 24(2)), provide secure IT-solutions (Article 25) and have research advisors and legal staff provide guidance when needed. However, the final hands-on work to ensure GDPR-compliance for a specific research project is done by the PI. As a result, **many researchers mistakenly see themselves as the controller and as legally responsible for the data processing**.

I suspect that this creates some unwanted consequences, particularly when it comes to handling potential data breaches. Most researchers working with personal data do not have a legal background. It can therefore be overwhelming to understand and interpret what GDPR requires upon a data breach. Even though there are legal personnel and a DPO at the institution that advice on the interpretation, it can be unclear who will actually do the work of crafting the data breach notice to the DPA or contact all research participants to inform of the data breach[16]. Since the PI have been responsible for all controller-tasks related to their particular project so far, it would be natural for the PI to assume they largely would be left with the work and responsibility of handling the data breach. For researchers who may be already overwhelmed, overworked and perhaps already are delayed in their project, it will in practice create a high threshold to report a data breach.

There is a need for guidance of who and what a controller is that is more adapted to the culture of academia and the principles of scientific research. It could be beneficial to specifically recommend that other personnel at the institution take over some of the hands-on work, such as determining the legal basis (Article 6 and 9) and deciding on whether a DPIA is required (and filling in part of the DPIA). This will hopefully create a stronger signal and experience that it is indeed the University/Institution, which is ultimately responsible as a controller, not the individual researcher.

As a further note on **Paragraph 137**: Although it may be true in a strict legal definition that PIs "are acting on the authority of their university", believe me, that formulation will most likely not resonate with most academic researchers. The university provides a framework for the research to happen and have to approve the research project, but the PI (the researcher) is leading the project. Any university has a duty to respect and uphold the

[16] In many cases researchers only have a signed piece of paper from the participant with no contact information. It will be very time consuming to track down and find contact information.

academic freedom of its academic staff and will only interfere and take authority if crucial policies, laws or ethical standards are violated.

7. More reflection on the **risks for scientific integrity** by uncritically promoting anonymization, deletion and data minimization

The guidelines rightfully list Verifiability as one of the key factors and criterion of scientific research (**section 2.1, paragraph 11 iii**). It states: "The research activities aim to achieve verifiable results and the conduct of research allows the hypotheses, methods, data and conclusions to be open to criticism" (my underlining).

Verifiability is often impacted by deletion, anonymization and sometimes data minimization. Anonymization can in many cases remove the chance to verify the validity of the research. Such as:

- The only way to verify qualitative (interview-based) research may be to contact the research participants themselves or to assess directly identifying video or audio recordings.
- In medical research, one often would need to check the patient's journals to verify the data.

The EDPB and other privacy experts appear to promote anonymization as soon as possible or that identifiable information is no longer needed in the research project. This may be as soon the results are published or as soon as the data can be pre-processed for data analysis and identifying information is not necessary for the statistical analysis. However, even if the researcher(s) themselves can fulfill the research project with the anonymized data, there is still an important need for society to be able to verify research results.

Ask yourself, would the *Lancet* MMR vaccine-autism fraud scandal have been uncovered today if Andrew Wakefield had followed the EDPB recommendations on storage limitation and to fully anonymize as soon as possible?[17] (No, it would not)

This is why many countries do have mandated archiving of research data in the law. In Norway this is only mandated within Health research[18], but I would argue that is important to be able to verify all scientific research results. Unfortunately, it is the case

[17] The investigation into validity of the research data started 9 years after the scientific publication. Fraud was finally confirmed as the research data did not match hospital records of the participants. See: *Flaherty DK. The vaccine-autism connection: a public health crisis caused by unethical medical practices and fraudulent science. Annals of Pharmacotherapy. 2011 Oct;45(10):1302-4.*

[18] See Health Research Act (LOV-2008-06-20-44), as amended by LOV-2025-06-20-71, new § 12 a, once in force. Data necessary to verify results must be stored for at least 5 years. The ethical committee can mandate longer storage.

that the rights and freedoms of data subjects are in direct conflict with the needs for ensuring scientific integrity. The EDPB needs to reflect more on these two conflicting societal needs, and the revised guidelines should address this. Retention of research data for verification purposes should not be only recommended or allowed where the national law happen to explicitly demand it, it should be applied for all research in Europe. Research is very much an international endeavor and results produced by one country will benefit the whole world. Similarly, impaired scientific integrity and fraudulent publications can hurt the whole world. I can only hope that the new European Area Act will eventually legally enforce long-time storage of research data, including personal data. But until then, **it would be very beneficial to the scientific community if the EDPB could recommend archiving research data** (with Article 89 safeguards) **and show more caution regarding anonymization and deletion**.[19] The EDPB should keep in mind that any "researcher" with low scientific integrity who wants to keep the option of committing research fraud, will more than gladly follow the EDPBs guidelines to anonymize/delete as soon as the data has been 'adjusted' to give the results they want.

[19] The EDPB (and EDPS) has a lot of power on how GDPR is applied in the European scientific community. It does not help that other EU initiatives or authorities, such as ERC and EOSC, try to promote archiving of research data, when EDPB and EDPS recommend (enforce) an a GDPR-interpretation that is incompatible with ensuring high scientific integrity across Europe.