Date: 24 June 2026

Subject: comments on the Guidelines 1/2026 on processing of personal data for scientific research purposes

With this input, UMCNL, COmmissie REgelgeving ONderzoek (COREON), Nederlandse Vereniging voor Pathologie (NVVP) and Health-RI jointly emphasize the importance of sound and practically workable Guidelines 1/2026 on processing of personal data for scientific research purposes.

We would like to share a few points of attention in response to the internet consultation on the EDPB Guidelines 1/2026 on processing of personal data for scientific research purposes.

In general, we welcome that the EDPB recognizes the importance of data driven research to improve (amongst other things) health care and innovation. Further, we welcome the additional guidance given by these guidelines to this field.

Despite of its best efforts, we feel the guidelines may be improved to foster health data-driven research. First, we summarize our main concerns. In the second section, we explicate these and some minor points.

Our main concerns are:

Definition of scientific research
The definition of scientific research may need some refinement. Firstly, it is unclear whether hypothesis generating research is given sufficient room. Secondly, the same question applies to applied research. Both serve extremely important roles in acquiring essential new knowledge and applying it in real life. In general, alignment with the demarcation of scientific research in the EHDS is essential for health data driven research.

Alignment with data driven research practice
There seems to be hiatuses in the recognition of how present data driven research works.

Important concepts that are essential to data-driven research and often play a role in adhering to the GDPR, such as FAIR, the data research life cycle and the DMP (data management plan), are not mentioned and taken into account. Both are required by funders and especially the first has consequences for GDPR compliance. More specifically, the FAIR principles show the tension between data-driven progress and individual rights, and

illustrates the sought balance between these under the GDPR. The DMP is more technical but is nevertheless a cornerstone in the curation of research data. It has a clear link with article 5.1.e GDPR and more. The DMP is an element of the 'research data life cycle', which as such is not mentioned either though the sentence above section 2.3 (at p. 13) comes near. We understand that the draft Guidelines follow the GDPR clauses more than the data research life cycle, yet, this exemplifies the gap which still, in spite of its best efforts, exist between the EDPB approach to data driven research and the research community.

Related to this, the role of research data infrastructures, and the challenges they face in their function when facing personal data (e.g. supporting multiple purposes with associated pseudonymisations and disclosures, FAIR data preparations at significant scale), are underrepresented in the guideline.

Moreover, the differences between research data infrastructures, data repositories and databases and their respective purposes are not reflected in the guideline. Here as well, alignment with the EHDS is essential.

Recontacting and consent

We agree that transparency about research is essential. Nonetheless, individual recontacting may not always be feasible or desirable. The bar for when recontacting would be disproportionate is set higher than we feel may always be feasible or desirable in (health) data driven research. Also here, the EHDS also states how individual recontacting should take place, which should be taken into account in the guidelines.

Similarly, research participants may suffer from so-called 'consent fatigue', making them less likely to re-consent or provide additional consent through dynamic consent, even though their willingness to participate in research has not diminished. A similar effect ('information fatigue') may occur when data sharing though research infrastructures is stimulated through e.g. the EHDS. Opt-out may be a better fit in certain situations than the proposed dynamic opt-in. Also, the EHDS makes an opt-out mandatory in member states for secondary use. Alignment between the opt-out requirements under the EHDS and these guidelines is necessary. National public opt-out registries and transparency portals with adequate public information – and communication campaigns is required.

EDPS-SRB decision

We feel the EDPS-SRB decision is relevant in many situations where data are provided for scientific research. In many cases, data may not be personal data from the point of view of the receiver. The reference to the 2015 Opinion on anonymization should be removed, as this opinion was refuted in the Breyer case. Further, we feel that article 11 GDPR may be relevant in situations in which the data are still personal data, which has not been given the relevant attention in the guideline.

We feel the part on genetic data and anonymity is explained too restrictively. In practice, a researcher may only use parts of an individual's genome, making re-identification very unlikely. Moreover, the context in which the genetic data are processed is relevant as well, e.g. when genetic data are processed in an SPE and researchers are forbidden to re-identify the individual, as will be the case under the EHDS. Classifying specific data types into an either-or category is undesirable and practically not feasible. Rather, institutional review

boards can play an important role in determining the relevant data protection measures based on the specific data used in research.

Alignment with the European Health Data Space (EHDS)
Research on human health data is extremely important to further future health and health care. In the near future, the European Health Data Space (EHDS) will guide health data driven research to a large extent. Therefore, it is essential that the current guidelines are in line with the EHDS, as we already mentioned in a few specific cases above.

Given that the EHDS regulation will be extremely relevant for health data research in the near future, more attention may be given to how the GDPR and EHDS should be interpreted in context, especially regarding legal bases. Further, more guidance is needed here regarding the EDPS-SRB decision and the context of the EHDS in which data protection is highly regulated, e.g. through the use of an SPE. Clarity on the interplay between relevant legislation is essential for health data research.

Detailed feedback

1. Description of research

We understand that the EDPB does not attempt at section 11 to formulate one encompassing definition of scientific research but mentions and elaborates on 4 constitutive elements. Elements i, iii till vi clearly define scientific research as distinguished from other social activities.[1] That being said, we wonder whether element i leaves sufficient room for hypothesis generating research. Such research is not merely 'exploratory or in a preliminary stage'. Especially exposome research is an example, where huge amounts of data from various sources need to be processed.[2]

Element ii adds a strong normative aspect which in general should be avoided in definitions or a description of the object to be regulated. The clauses which regulate the normative aspects come later in the substantive clauses of the regulation or regulations. Yet, given that those are largely absent in the GDPR itself, we acknowledge that some normative element can be embedded in the description to distinguish 'good' from 'bad' research. Though the 'concept of consent to participate in research' should not be a constitutive element. It would mean that much public health research, which is often not based on consent, would not be research anymore. Also, in the near future, the EHDS provides much room for scientific research not based on consent. This also applies when taking into account that the EDPB does not necessarily refer to GDPR consent in this context but, as stated later in the draft Guidelines to some form of 'ethical consent'. For much public health legislation in EU countries based on articles 9.2.i and j GDPR also 'ethical consent' does not apply for the

[1] See for the distinguishing factors H. Collins, R. Evans, Why democracies need science, Polity Press, Cambridge, 2017, especially p. 55.

[2] E.g. Safarlou, C. W., M. Van Smeden, R. Vermeulen, and K. R. Jongsma. 'Enabling the Nonhypothesis-Driven Approach: On Data Minimalization, Bias, and the Integration of Data Science in Medical Research and Practice'. The American Journal of Bioethics 23, no. 9 (2023): 72–76 with further references. https://doi.org/10.1080/15265161.2023.2237452.

collection of the data and/or the research based health care data.[3] The reference to ethical guidelines etc. which indeed usually mention 'consent' as the default, is somewhat misleading or even biased as these all allow for an exception to the consent principle. The Allea CoC on research integrity (ECCRI) does indeed mention (training in) ethics but not consent.

Additionally, we wonder why elements v and vi come last. The earlier elements describe how research should be performed, elements v and vi describe why research is performed, the essence of it. Hence, why it is essential for an open society and why the freedom of research has a special status in the Charter of the EU. We recommend reordering the elements, so that this essence comes first. Importantly, element v should be brought in line with the definitions on scientific research under the EHDS. Under the EHDS, scientific research is a rather broad category, including algorithm development and innovation. The Guidelines should encompass these as well. Processing of health data under the EHDS should still follow GDPR regulations. Any discrepancies between these Guidelines and the EHDS should be avoided, as they will lead to an unworkable situation for anyone involved in the secondary use of health data. Further, it should be made explicit that applied research falls under this category as well.

Lastly, we strongly recommend deleting in element v: 'and well-being'. This phrase seems to come from the definition of research in the Digital Omnibus Act but cannot be found in any other major statements about the essence of scientific research, except -perhaps- in certain countries with autocratic regimes. The phrase opens the door to censorship. Research and its outcomes can also be confrontational and may not contribute to society's (who is that society by the way)[4] wellbeing on the short run.[5]

2. Presumption of compatibility (par. 3.1)

We very much welcome the clarification of clause 5.1.b last sentence. In principle 5.1.b supersedes 6.4 GDPR. Section 3.1.2 is also very helpful in this context: if further processing for research is done after a transfer, and also the transfer by the original controller is based on 5.1.b, the new controller does not need to undertake the compatibility test of 6.4 either. Of course, the new controller does need a legal basis and, in our field, also compliance with 9.2 GDPR and national implementing legislation.

The EDPB clearly follows the importance given by the legislator to 5.1.b, last sentence, implying that further processing for statistics and scientific research is inherent to the

[3] That legislation does not always mention scientific research as such, but certain purposes for the collection and analyses of health data. Just as the allowed purposes under the EHDS mention research as (only) of them. But none of these other purposes can be reached without scientific research according to the Allea ECCRI definition. So, this is all scientific research even though the purposes for which this research may be performed are more narrowly described.

[4] When mentioning the knowledge of society, this 'whole of society' plays a lesser role. Knowledge can spread through society unevenly and generally is. When speaking of wellbeing of society, there is an implicit notion of society as a whole. Yet, who defines that and certainly what constitutes its wellbeing?

[5] A clear example is the present curbing of NIH funds in the USA for research which investigates the relation between health, gender- or racial inequality. Such research is said 'to promote dangerous ideologies which undermine public health'. NRC 13-14 June 2026, W 16.

original purpose. In that context we were surprised by point 84 of the draft guidelines, which states

> If a controller that collected personal data for another purpose than scientific research at a later point in time decides to further process the personal data for scientific research purposes, then the controller must inform the data subjects accordingly (nt). In line with the principle of purpose limitation, this also applies if the controller intends to process the personal data for a new scientific research purpose that is different from an initial scientific research purpose (nt)

Point 84 is clearly incompatible with the intention of the legislator and section 3.2 of the draft guidelines. We will come back to the much less nuanced, compared to the rest draft guidelines, section on transparency at section 6 of this comment. Suffice for the moment that point 84 should be deleted or very much nuanced.

The EDPB makes an exception for when the legal basis is a legal obligation or (GDPR) consent. We do not understand the legal obligation case. There is long tradition for entities processing personal data based on a legal obligation also use those for statistics and research.[6] We understand the limitation better when the legal basis is GDPR consent. If there are clear promises for what kind of research the data are being used, those promises should in principle not be broken. Yet, for example with an original narrow consent (as may have happened originally after the advent of the GDPR and the guidelines on consent which did not seem to allow for broad consent as is the case in the present draft), a time lapse and new important opportunities for research on those data, we believe that research on those data should be possible if certain conditions are met. For health data, this would also correspond to the regulations under the EHDS. We will come back to this issue on three occasions: the discussion on the legal basis and GDPR consent, and recontacting participants.

To summarise that discussion, a three-tiered solution should be available here.

a. To further analyse those data can be based on 5.1.b last sentence;
b. The controller has another legal basis (such as 6.1.e) for the envisaged research;
c. An ethics committee or similar committee, such as a data access committee has decided that these data may be used for the new research, based on criteria that reconsent or opt-out is not reasonably feasible and that the new research is not clearly incompatible with the original consent.

Especially the a>b is important here. Based on the earlier EDPB Guidelines on consent, it has been believed by many that once based on GDPR consent, the controller cannot ever use another legal basis anymore (‘one cannot change horses during the race’). This has had a stifling effect on health research with older data repositories based on rather narrow GDPR compatible consent. The present draft guidelines seem to imply that one can, if certain conditions are met. We endorse that reading. However, a clarification in this context would be helpful.

On another note, we feel some more guidance is needed here regarding paragraph 21: At first glance there seems to be a direct contradiction between paragraph 20 citing Recital 50

[6] Example of banks in anti-money laundering obligations. Public health institutions handling personal data based on the WHO Treaty on infectious diseases.

GDPR "should be considered to be lawful *operations*" and paragraph 21 talking about this not necessarily satisfying the "*principle* of lawfulness". If the intention is to distinguish between the "principle of lawfulness" Art 5.1(a) and the "lawful base" Art 6, this could be made more explicit in the guidance? And could it be clarified to what obligation / assessment this specifically leads? How are the lawful base and principle of lawfulness determined separately?

4. Storage limitation (par. 3.2)

We welcome the flexibility which seems to be offered by this section. However, the following sentence at point 27 might give rise to misunderstandings: 'the storage period must be determined by the controller before the personal data are processed'.[7] The rest of the paragraph clarifies that a fixed period often cannot be determined beforehand or when that has been the case, may be prolonged. Researchers organise these issues in the DMP, including periodic review whether it is still necessary to maintain the data in their original form.

5. Consent (par. 4.1)

This paragraph is a great step forward in recognising the inherent complexities of longitudinal research projects, especially in the life sciences. Broad consent on which many large cohort studies with volunteers are based, is now explicitly recognised. In the life sciences the consent, when that is the legal basis, must also be 'explicit' (9.2.a GDPR). We appreciate that the explanation of explicit consent at point 67 does not set the bar higher in this respect. We recognise that some specification must be given, though the overall aim of the cohort may still be very broad, such as 'healthy aging'. Such broad descriptions have never deterred many volunteers to participate or ethics committees to approve such projects. Indeed, additional 'safeguards' must be set in place. Such GDPR safeguards go hand in hand with the researchers' interest in retaining the interest and involvement of their participants for long term follow-up and possible new waves of questions. Yet, retention studies[8] show that the interest often wanes, despite best efforts to retain attention and make participation interesting or even fun. 'Participation (meaning active participation) fatigue' go hand in hand here with the better-known consent fatigue. Seeing participants as wilful agents should strike a balance and respect their choices, also if that means not responding anymore. Not responding anymore does not mean withdrawing consent. This has consequences for the section on 'dynamic consent'. With such consent mechanisms one never knows whether the non-responding participants don't agree or simply are not interested anymore. We propose

[7] The EDPB refers to article 25.1 GDPR in this context. However, we wonder whether this clause should be read in this specific way. The article requires thinking of data protection by design and default. That can also mean a mechanism to decide about the ultimate storage period at a later stage.

[8] Teague, Samantha, George J. Youssef, Jacqui A. Macdonald, et al. 'Retention Strategies in Longitudinal Cohort Studies: A Systematic Review and Meta-Analysis'. BMC Medical Research Methodology 18, no. 1 (2018): 151. https://doi.org/10.1186/s12874-018-0586-7; Teixeira, Raquel, Julia Doetsch, Ana Isabel Freitas, et al. 'Survey of Data Collection Methods and Retention Strategies in European Birth Cohorts of Children and Adults Born Very Preterm'. Paediatric and Perinatal Epidemiology 36, no. 5 (2022): 706–14. https://doi.org/10.1111/ppe.12845.

that 'assent' is a better solution or at least a viable alternative. This means that the participant gets a notification about new tests or a direction of the research which are not within the reasonable expectations. If the participant does not agree that her data will be used for this new direction of the research, she can object. Of course, the notification should clarify that without a response, the data will be used for the new direction of the research, in short: a spin-off study.[9] As a – common sense – rule of thumb, if people disagree, they are much more inclined to let you know their disagreement than when they agree.[10] Therefore it may reasonably be expected that those who disagree will not be included in the new spin-off study. Such an approach would prevent the inclusion bias[11], and hence possibly incorrect or inclusive results.

The results of this new spin-off study will be based on data which have been acquired under GDPR consent. Yet, the spin-off study cannot (solely) be based on GDPR consent. This underscores the pressing need for recognising that during the research data life cycle, the legal basis can change. The Guidelines on consent seemed to deny this option. The stifling effect of this position for older repositories has been mentioned already. This position could have a stifling effect also here, leading to research waste, meaning that valuable data cannot be used anymore due to a strict interpretation of the GDPR which also the average citizen wouldn't understand. They have had their say, didn't object but the data to which they originally consented and often actively contributed[12], can nevertheless not be used because the legal basis cannot change. We note this for clarities sake as implicitly the EDPB recognises that the legal basis can change, example given at point 108, second dot. In the example given above, with a spin-off of the research, that information could be given in the context of the announcement of the spin-off study with the possibility to object.

The possibility to object, or opt-out, will become more common to citizens with the upcoming EHDS. Secondary use of e.g. cohort data under this regime will be governed by an opt-out mechanism. Posing the possibility to opt-out therefore adds to the understandability of citizens of health data use for scientific research. It is essential to them that regulations for scientific research are uniform and clear. And, it prevents citizens from developing consent fatigue. Moreover, uniform regulations and essential to researchers and research institutions as well. In order for them to be able to efficiently and effectively use health data to further knowledge about health and disease and make progress in health care, clear and uniform regulations are necessary.

---

[9] As this study will still be primarily based on the data collected already.

[10] In the psychological literature this known as the 'negativity bias'. (AI generated answer to a prompt)

[11] About bias via consent systems see: De Man, Yvonne, Yvonne Wieland-Jorna, Bart Torensma, et al. 'Opt-In and Opt-Out Consent Procedures for the Reuse of Routinely Recorded Health Data in Scientific Research and Their Consequences for Consent Rate and Consent Bias: Systematic Review'. Journal of Medical Internet Research 25 (February 2023): e42131. https://doi.org/10.2196/42131; Hermus, Merel, Celinde H. Scharloo-Karels, M. Arfan Ikram, et al. 'Opt-In versus Opt-out for the Secondary Use of Routinely Recorded Health Data: A Randomized Controlled Trial'. European Journal of Internal Medicine 133 (March 2025): 100–105. https://doi.org/10.1016/j.ejim.2025.01.017

[12] Via questionnaires etc.

We fully agree that there should always be an option to withdraw consent for the data processing via an active notification to the controller and that participants should be aware of that option. However, footnote 70 of the draft guidelines describing the consequences of withdrawal is not fully correct. It misses that the data may be kept under article 17.3.d GDPR. The draft guidelines discuss that research exemption and we will come back to it as well, but not mentioning 17.3.d in footnote 70 is an omission and may create misunderstandings.

6. Obligations to inform, especially when data are collected directly from the data subject (par. 5.1-5.3)

Obviously, we agree that research should be performed in a transparent way. Many researchers make great efforts in this context, though not in the formalised way the EDPB sees transparency. They interact with patient organisations, present at their meetings, organise the (annual) meetings of the cohorts with volunteers. And indeed, these efforts might not reach the average younger and middle-aged citizens who occasionally visit health care for lesser problems and are largely unaware that the health care they receive now, is based on earlier research. Health care and research have certainly with the advent of empirical methods in the 19th century always worked in tandem.[13] As it called nowadays 'real world data' or 'learning health care systems' have been given explicit recognition in the last decade but were there on the background. We should make a distinction at this point:

- Data explicitly collected from volunteers: the, usually long term, cohorts we referred to earlier when discussing broad consent.
- Secondary use of data of personal data collected during health care delivery (diagnostic procedures, care given and their outcomes). This also falls within article 13 GDPR at least at its initial stage,

In the previous section on broad consent, we mainly focussed on the first example. It should be mentioned that to be valuable for the research, these data will always be combined with data from other sources, such as patient records, registries and data at, in the Netherlands, Statistics Netherlands.[14] The participants give consent to enriching their data with data from these other sources. We generally agree with what the draft-Guidelines state about transparency and interacting with the participant in this context. We underline, however, that certain trade-offs between utter transparency and over burdening the participants will be needed here. We have mentioned already that participants may lose interest. And they may lose interest more easily, in the sense of 'don't bother me with all this, I have given consent already', when being overloaded with information.

The second example of collecting from the data subjects is the focus of this section. We challenge the position of the EDPB on this point. At section 3 of our comment we mentioned that when taking the EDPB interpretation of 5.1.b last sentence seriously, an interpretation which we endorse, point 84 of the draft-Guidelines is in clear contradiction. 'Scientific research is inherent to the original purpose and can in principle based on the original legal basis, but yet, further use for scientific research should -as it seems- for each new research

[13] Roy Porter, The greatest gift to mankind, medical history of humanity from antiquity to the present, Fontana Press, London, 1999, especially Chapter XVIII.
[14] Example 8 is also in that sense not aligned with the realities of health research.

project be nevertheless explicitly notified to the data subjects.' Legally both positions are incompatible. And let's contrast the latter position with reality. In (Dutch) health care the primary collection of patient data is not based on consent. The professional is obliged to keep an accurate medical file. Secondary use for research for a learning health care system is in a certain way the default position. Certainly at larger health care providers patient data are always used for observational research at some point in time. The researchers might be:

a. those embedded in their health care team;
b. others at the health care provider;
c. recipients outside this controller, or;
d. a combination of a/b and c.

In case a there is not an issue of medical confidentiality which always comes first. Neither is there an issue with GDPR purpose limitation, as argued. This aligns with a long tradition in health care, reflected -also in the newest version - of the Hippocratic oath[15] with the line: 'I will advance the medical knowledge of myself and others'. Though many patients and probably also the EDPB might be unaware of this, present health care wouldn't exist without this responsibility of doctors. Of course, there should be transparency. Privacy notices of Dutch health care providers usually clarify that patient data may be used for research and when this involves access to directly identifiable data by others than their health care team, this will be based on consent. Yet, others will hardly ever get access to directly identifiable data, and then a different regime applies to which we will come back.

In case b there might be an issue with medical confidentiality. Certain legal systems like Germany, however, consider the whole health care provider as the treating entity and allows research with the patient data.[16] The Netherlands is stricter on this issue. In principle consent should be asked. However, as this secondary use for research is so preponderant, potentially involving all patients, also the ambulant patients for a broad array of research questions, usually the opt-out system as allowed by Dutch law, is being used.[17] From a GDPR perspective, the researchers at the health care provider will only get access to pseudonymised data, via a digital research environment within the health care provider.

In situation c there will be a potential third party recipient in the sense of article 4.10 GDPR, if the data are not pseudonymised-anonymous. The privacy notices should clarify that. There are many such 'recipients'.[18] The draft guidelines treat research as rather isolated events, but

---

[15] Which can be downloaded – in Dutch – via: https://www.knmg.nl/actueel/publicaties/publicaties-per-onderwerp/artseneed

[16] See the Chapter on Germany in: Veen, E. B. van, and R. A. Verheij. The Rules and Procedures in The Netherlands Compared to Denmark, England, Finland, France, and Germany. MLCF ; NIVEL, 2023. The report can be downloaded from: https://www.nivel.nl/nl/publicatie/further-use-data-and-tissue-learning-health-system-rules-and-procedures-netherlands .

[17] See the Chapter on the Netherlands in: Veen, E. B. van, and R. A. Verheij. Further Use of Data and Tissue for a Learning Health System: The Rules and Procedures in The Netherlands Compared to Denmark, England, Finland, France, and Germany. MLCF ; NIVEL, 2023.

[18] Apart from to Statistics Netherlands (sometimes via a tiered system with Dutch Hospital Data as a processor in between, and for primary care a representative sample via Nivel Primary Care Database care registry), certain disease registries, quality registries (which can also be used for research). About the Nivel Primary Care

as mentioned, health care and secondary use for research work in tandem. The privacy notices will describe categories of 'recipients', giving some examples. At least article 11 GDPR will apply to those recipients outside the controller (= health care provider). These data might also be pseudonymised-anonymous data and then those formally not recipients in the sense of the GDPR anymore. We will come back to this at section 7 of this comment.

In all the cases a-c (d) preparing the dataset is secondary use but in line with the 5.1.b GDPR last sentence. The privacy notices in the sense of article 13 are general. The right to object according to Dutch law is achieved by general patient leaflets in conjunction with other important information about the visit to the health care provider. It can often also be found when patients log in to their personal site at the health care provider.

6. Information when personal data are not collected directly from the data subject (par. 5.4 of the draft Guidelines)

A close reading and comments at many sentences would take too much space. In general, though we agree with the general concept of transparency also at controllers who host research databases in the life sciences, we submit that this section is based on misunderstandings about:

a. what kind of data reach these controllers,
b. their options even if they would have contact addresses,

and more fundamentally,

c. the essence of their activities and their alignment with the purposes of secondary use of patient data for research.

Ad a:

We have mentioned already that the data always reach the third-party recipient in a pseudonymised way. The draft-Guidelines fail to mention article 11 GDPR and we can only wonder why. We consider it a serious omission and submit that if the data are not already pseudonymised-anonymised (see the next section of this comment), the processing at the recipient with research database, at least fulfils the criteria of article 11. The rebuttal might be that the original controller could give the pseudonym to the data subject. In which case, by the way, article 11 would still be applicable but an article 11.2 case might arise. Sophisticated pseudonymisation does not work like this. The controller does not know the final pseudonym. Additionally, even if the controller would know this, some patients would be overloaded with pseudonyms, and the ICT systems of health care providers with managing sending the correct pseudonym to the correct patients, as well. Resources which in our opinion could be better used for the primary function of health care.

Ad b:

Database see: Van Hommerig, Joost W., Robert A. Verheij, Karin Hek, et al. 'Data Resource Profile: Nivel Primary Care Database (Nivel-PCD), The Netherlands'. International Journal of Epidemiology 54, no. 2 (2025): dyaf017. https://doi.org/10.1093/ije/dyaf017

As argued, the new controllers don't have the contact addresses in the case of secondary use. Article 11 resolves the 'catch 22 situation, that in order to comply with the GDPR the controller should retrieve identifiable (contact) information while the essence of the 'privacy enhancing technologies' is that the controller shouldn't have those. The latter takes preponderance. But now, supposedly, that these controllers have some additional information. In that context we submit that the draft-guidelines require efforts of controller which are usually not feasible or drain the research budget. Additionally, the draft-Guidelines underestimate the 'bias problem' (see footnote 11) which does not only apply to 'covert research' as discusses in the draft Guidelines.

Ad c

The EDPB states at point 103 of the draft Guidelines (our emphasis):

> ….processing of personal for scientific research without *individually* informing the data subjects should only occur *where it is strictly necessary*, as it constitutes a *serious* interference with the fundamental rights and freedoms of data subjects.

We have deep concerns with this sentence. This sentence juxtaposes research and individual rights as utter opposites, with the latter having complete preponderance. Both research and data protection have found recognition in the Charter of the EU. Unlike this sentence, other parts of the draft Guidelines recognises the balance which needs to be struck here. The quoted sentence is unbalanced, reflecting in a way the EP approach to research under the GDPR by rapporteur Albrecht, in the EP's version of the draft GDPR. That approach has been refuted in the final text, amongst other things by the then 'data save lives' initiative of the Wellcome Trust. We have mentioned already that health care and research via secondary use of patient data work in tandem. Third-party controllers with research databases where data from various health care providers are being combined, as the patient must be followed over time and through its trajectory in the health care system, are an essential part of this ecosystem.

The second concern is that it mistakes the reality of the pseudonymised data chain. We discussed that already. But the quoted sentence means in essence that also when it is not strictly necessary, in that limited way as the EDPB sees it (with its exception for 'covert research') but unfeasible or not possible (see ad a and b supra), the researchers seriously interfere with fundamental rights. Though they do have a legal basis to process these data, assuming that there are still personal data.

Lastly, this sentence will rub health researchers and the ecosystem in which they are embedded, the wrong way, hence also the members of privacy committees or data access committees who approve or supervise the specific projects executed with the data of the repository and to their DPO's. While dedicating their career on improving health care and patients'/citizens safety, they are considered as seriously interfering with the fundamental rights and freedoms of patients from whom those data originate and, when not too late, might also profit from that research.

In the same vein, this part of the draft Guidelines seem to contradict the regulations for secondary use of health data under the EHDS. When – under the coming EHDS - researchers analyse data in an SPE after a data permit, data subjects are not individually informed about this processing. Yet, the EHDS prescribes another procedure to provide transparency and certainly strikes a balance between an individual's fundamental rights and societal value. The

draft Guidelines however, seem to imply that this process, with all of its guarantees, is seriously interfering with fundamental rights. Any such discrepancies between the Guidelines and EHDS should be avoided.

The mentioned sentence should be deleted or very much nuanced.

7. Pseudonymised data (par. 8.3 of the draft Guidelines)

Though discussed later in the draft-Guidelines we raise the issue now, following up on the discussion on what kind of data are received by third-party recipients. We note that the EDPB refers to the draft guidelines on pseudonymisation, even though they were issued before the EDPS/SRB case and should be seriously reconsidered after the judgment of Court of Justice of the EU. We welcome that the EDPB mentions this judgement in the present draft. However, the draft Guidelines hardly reflects the consequences. In the previous sections we concluded that draft Guidelines miss some: the third party 'recipient' will often receive pseudonymised-anonymous data. The Guidelines should not refer to Opinion 4/2015 of the then Article 29 Working Party anymore. Many authors considered that Opinion outdated after the Breyer judgement and it's certainly not a valid interpretation of the law anymore after mentioned EDPS/SRB judgment. Is is essential not to diminish the consequences of that judgment, for example by the recognition of pseudonymised-anonymous data in the data chain to research databases.[19] Especially when in the near future health data are processed under the EHDS, where data users get access to data in an SPE, the consequences of this judgement cannot be omitted, as they will have serious consequences for determining whether data should be considered personal data from the point of view of the data user or not.

At point 9 we will discuss how in the pseudonymised data chain data subjects' rights can be preserved even though the third party 'recipient' is unaware of the identity of the data subjects.

8. Rights of data subjects (par. 6 of the draft Guidelines)

At point 118: We generally agree that the controller must substantiate why the exemption to the right to erasure applies in each specific instance. But these are not ad hoc decisions. The controller of the research database should clarify the general rules on which these decisions are based and why these rules are there, as an exception to the general regime of the GDPR. This increases transparency but also the general understanding of the public.

We miss that one of the reasons not to delete the data being to substantiate the integrity of the outcomes of research, hence after the research has ended, when deleting datapoints create a substantial risk of distortion of the results that cannot be corrected and would make the results unreliable.

Ad point 124: This point makes mention of possible Union of MS law that may 'compliment' the provisions of article 21.6. It would be fairer to state that MS may derogate from article 21.6 via national legislation based on article 89.2 jo. article 9.2. i or j GDPR. The Dutch legislator didn't apply this exemption for research in general, but did so in specific legislation.

[19] As mentioned, if not pseudonymised-anonymous, then article 11 will apply to those controllers.

So, outside those exceptions article 21.6 applies. Yet, this clause is – in the Dutch context – less relevant than the general opt-out or right to object on which most secondary use of health data for research is based. That opt-out is by itself and derogation of article 9.2.a GDPR based on Dutch MS law implementing article 9.2. i and/or j.

Ad example 19: This example may lead to discussion, because many readers may conclude that updating the name of the data subject in the repository would be a reasonable solution as well. We suggest that a different example is chosen where such an alternative is not possible; alternatively, the example could clarify why changing the name is not an option.

9. Rights working through the data chain (not a specific point in the draft Guidelines)

Earlier it was argued that the draft Guidelines misunderstand or underestimate the pseudonymisation in the research data chain. The research database for secondary use in the case of a third party 'recipient' cannot comply with the demands of article 15 because of the PET's used in that data chain. Yet, we like to stress that this certainly does not mean that the rights of data subjects are meaningless. The Dutch Code of Conduct on health research[20] (Dutch CoC) stresses that the right to object should follow the data chain, as also follows from article 19 GDPR. The objection will be made at the first controller, the health care provider of the patient. When the patient has objected, the data will not be forwarded. If the patient would object at a later moment, the recipient will be notified via the pseudonym, which can be repeated. In that case article 17.3.d will apply to those data.

The Dutch CoC also stresses the importance of general transparency at the controllers of those research databases. They have extensive websites with information accessible to lay people. The data are carefully curated, data security follows ISO 27001/2 and NIS2. They are only opened-up[21] after the approval of a privacy -, data access or similar committee. Most of these committees have one or more representatives of the relevant patient organisation and researchers interact with patient organisations also in the context of the research agenda.

Hence, the whole secondary use research data ecosystem, is not an all (GDPR), or nothing at all, pseudonymised-anonymous data, issue. The ecosystem is a complex whole where various stakeholders interact to improve health care via patient data.

In general, much emphasis on informing data subjects even when contact details might be inaccurate, public registries should be used then. The bar for when recontacting would be disproportionate, is set very high.

This information is also very relevant for data infrastructure providers who may serve many different recipients, each with their own data processing purposes; the requirement to inform frequently could lead to *information fatigue*. Regarding paragraph 91: In the case of research infrastructures, new pseudonyms may be created for each disclosure of data to a new recipient. Especially when Health Data Access Bodies provide access to data to data users on a grand scale under the EHDS, data from certain categories of data subjects may be used very often. Providing this as information every time to the subjects could lead to

[20] Code-of-Conduct-for-Health-Research-2022.pdf

[21] As transfer of data hardly ever takes place. Data can be analysed at the safe environment of the resource which then acts as a SPE.

information fatigue, while providing very little benefit to the subject. It would be beneficial to research infrastructures if the guideline would provide clarity about the level of clustering/bundling of similar information that is permissible.

10. Specific remarks regarding research health data infrastructures, especially in the context of the EHDS

As explicated above, research health data infrastructures can be complex systems that exist to stimulate the re-use of health data to improve health care and our understanding of health and diseases. Especially with the implementation of the EHDS in the near future, the Guidelines should take these complexities and the interaction of the GDPR and EHDS into account. Currently, the draft Guidelines do not fully align with the broader infrastructure model of the EHDS, which also includes secure processing environments. This need for alignment is not new. In Joint Opinion 3/2022 on the proposal for a European Health Data Space, the EDPB and EDPS already stressed that the interaction between the EHDS, the GDPR and Member State law must be particularly clear. In the construction of such infrastructure, it is important to build on legal clarity, as it can be hard to change data governance aspects once international collaboration agreements are signed. Changes in legal interpretations can mean setbacks of several years. Wherever there is room for interpretation, the construction of personal (health) data infrastructure therefore needs to err on the side of caution; this can mean that scientific goals are harder to achieve than could be considered necessary. In this paragraph, we focus on points that are specifically relevant building such infrastructures, especially in the context of the EHDS. Some points may overlap with the points stated above.

- Regarding paragraph 13: This text equates “research data infrastructure” with “a research repository or database”. There are many other types of research data infrastructure, e.g. analysis software, suitable secure processing environments, data and ontology standards, etc. It would be better to be clear here that this text only applies to research data infrastructure in the form of a research repository or database. Importantly, Research repositories or databases often support more than only research purposes. It would be useful to clarify whether this guideline solely applies to repositories and databases that exclusively support research reuse of data and in how far it could be acceptable for infrastructure that supports policy development, quality monitoring or other purposes alongside scientific research.

  Research repositories and databases are different entities as they serve different purposes: reference *databases* strive for a completeness or representation and generally have personnel that curate data into a standard form. Research *repositories* tend to collect data from depositions; they put constraints on the shape and format of the data. Yet, the responsibility for the data preparation and inclusion is with the depositor. The differences in purpose and ways of working give rise to diverging controller/processor relations for the different involved parties for different steps in the realisation of the infrastructure.

- Regarding paragraph 100: When someone responsible for a data processing activity cannot identify the data subject, it is also impossible to contact the data subject to provide information about the processing. If, following ECLI:EU:C:2025:645, such processing does not fall under the GDPR because the data is anonymous to the one

responsible for the processing activity, the obligation to inform also does not exist. In cases, however, where the activity still is subject to the GDPR but the subjects cannot be identified by the controller, it would give clarity if an explicit exemption from the information obligation would be added here. Alternatively, in case the data was obtained through a controller-to-controller transfer, the information could be provided to the subjects via the original controller without revealing the identity of the subjects to the new controller. This would require cooperation of the original controller (potentially a research data infrastructure provider).

- Regarding paragraph 126: a research database or repository could require processing of data preparing it for efficient re-use (e.g. harmonisation and FAIRification of data). It is difficult to read from the current guidance whether this efficiency gain can be sufficient to classify this processing as *strictly necessary*. Note that facilitating efficiency of data re-use is an important goal of research data infrastructure and the upcoming EHDS. Also note that the efficiency gains targeted here will lead to *less* processing of personal data *in total* counted over multiple cases of reuse.

- Joint controllership is especially common in research infrastructures. We feel some additional explanations would be useful here.

  Regarding paragraph 140: case law seems to favour the interpretation of many relationships as joint controllership, because it favours the subject rights (e.g. C-604/22). Is that not in conflict with the guidance given here? It is often suggested that wider research data infrastructure, including those that develop and maintain research software, could classify as Joint Controllers for research projects, severely complicating that type of research infrastructure work.

  Regarding paragraph 140: ECLI:EU:C:2023:949 under paragraphs 44-46 indicates that joint controllership is not a consequence of a joint controllership agreement; how is this compatible with parties being declared joint controllers by union or MS law?

  Regarding paragraph 143: We are confused by the phrase “rely on the same legal basis”. Would the parties not each need their own legal basis? The legal bases could differ specifically in instances where one of the controllers carries out a legal obligation (e.g. under the EHDS). Does this paragraph suggest that the legal bases of the different controllers would refer to the same basis in Art 6.1?

- Regarding paragraph 164: The sentence “data subjects should not be given the impression that their personal data will be anonymised if the data in fact will be processed in a pseudonymised form” seems to predate ECLI:EU:C:2025:645 and is hard to apply within a context where pseudonymized data are personal to a certain controller but anonymous to another party to which they are disclosed. This situation will be especially common under the EHDS, where data users will get access to health data in an SPE, and have no realistic (legal) possibility to re-identify the data subjects.

- Regarding paragraph 165: This guidance seems to be hard to apply to genetic information that is made available for scientific research in an SPE, e.g. such as foreseen under the EHDS. Since it will be very unlikely that the recipient of genetic data under those circumstances will have the reasonable means to identify the subject from the genetic data, it could still be considered anonymous under such conditions. This can also be read from footnote 243, where re-identification based on genetic data requires linking that data to external resources, a process that is impossible when the data is provided to a re-user in an SPE.

Best wishes,

On behalf of

UMCNL

COmmissie REgelgeving ONderzoek (COREON)

Nederlandse Vereniging voor Pathologie (NNVP)

Health-RI

Should you have any questions regarding this, please contact us through Susanne Rebers, COREON (secretariaat@coreon.org).