GDPR and research [1]

A basic overview

1. Introduction

The scheme on the next pages gives an overview of clauses in the General Data Protection Regulation which are specifically relevant for (health) research. The scheme was drafted to give guidance to both researchers and lawyers who need to implement the GDPR (and national legislation) in the context of research with personal data.

There are various reasons for drafting this scheme:

- The Wellcome Trust drafted a scheme[2] but that is hardly used, partially because the scheme is also meant to guide GDPR implementing legislation at the member states level and is not completely comprehensive. Yet, we are indebted to this scheme.
- for purposes of the MLC Foundation. We give guidance to research and legislation and self-regulation to come. We should have the 'facts' straight then;
- I have seen overviews of the GDPR in the context of health research where the exemptions[3] for research were not mentioned systematically. Or where it was put forward the research exemptions can only be used if the research is in the public interest. As we will see, the relation is more complex[4].
- I have heard mentioning reading the GDPR for the specifics for health research as if this was rocket science. Though the GDPR is sometimes a difficult read (see also hereinafter), finding those exemptions is not that difficult. Admittedly, their interpretation is;

As all schemes it has certain limitations:

- The other provisions of the GDPR, which are not specifically dedicated to or relevant for research, are not mentioned. Of course, these are very important as well and can influence the application of the research exemptions. An example is the data protection impact assessment[5], which is always relevant for health research (but for many other types of processing of personal data as well) and the status of one way

[1] The research leading to this overview was funded by the H2020 RECAP Preterm project, grant no. No 733280

[2] https://wellcome.ac.uk/sites/default/files/new-data-protection-regulation-key-clauses-wellcome-jul16.pdf

[3] One might also read 'research privilege' here. Though that 'privilege' is very conditional.

[4] To state it bluntly: all research is in the public interest but some research is more in the public interest than other research. Research is in the public interest in general as we need data which meet the criteria of research validity to feed the public and scientific debate as 'objectively' as possible.And such research in general can under circumstances fall back on reseach exemptions. Certain research can specifically be considered to be in the public interest when that would follow from national or EU law (see also footnote 6).

[5] See the Guidelines on Data Protection Impact Assessment of the Article 29 Working Party, 4 April 2017, revised 4 October 2017.

pseudonymised data. On another occasion we will come back to the latter.

- This is not an explanation of the GDPR as a whole but the text of the GDPR should be read as a whole. By focussing on one clause one might lose insight in the context. We tried to accommodate this problem in the right column (Discussion) by referring to relevant other clauses or the Recitals for explanation of the clause in question. Yet, a balance had to be struck here between becoming too lengthy and the readability of the scheme.
- The GDPR is the result of a political compromise. There is an abundance of words and often some stapling of ideas expressed in those words which do not always seem immediately compatible.

To continue on the last dot. All legal texts involve a certain interpretation. Hence why 'fact' in the second bullet above was between quotation marks. The dilemma of interpretation is especially the case with the GDPR which had a lengthy procedure before being finalised[6]. The draft text of the European Parliament (EP) amending the proposal of the European Commission (EC), discarded many research exemptions of the text proposed by the EC. Commentators, including patient organisations, pointed at the negative consequences for health research. In addition to practical problems of the text by the EP, these comments also stressed the common good through results of research to improve health, which would be jeopardised with the EP text. That EP text was basically about individual control. The final compromise after negotiations with the Council, brought back a balance. Though at the cost of leaving certain details to national legislation, especially regarding the 'exemption' for informed consent in the context of research with personal data.

The tension between individual control, or 'autonomy' in the bioethical debate or informational self-determination in the privacy debate, and the common good which data can also serve, cannot be discussed here. As facts are not simple facts when working with a legal text, it should be mentioned this author has a bias for the common good approach. At the same time that also poses obligations on research with data. Insofar as relevant for an overview, these will be mentioned.

As last remarks, how to use the scheme. This scheme is published as a form a 'creative commons' with the following restriction. When quoting part of it, or cherry-picking from it, reference to the full text should be made as it is available at the site of MedLawconsult and the MLC Foundation. As everyone can have access to the GDPR, texts of the clauses themselves are summarised. The proviso that, though as factual as possible, the scheme involves some interpretation and reflects the full text of the GDPR only partially has been said already.

[6] Coppen R. et.al. 'Will the trilogue on the EU Data Protection Regulation recognize the importance of health research?'. European Journal of Public Health (2015) Vol 25 no 5 757-758.

2. Scheme

What	Where	Specifics for research	Discussion
Research	Recital 159 (partially)	…	Research is not defined in the GDPR. Neither is 'public interest' by the way[7]. Recital 159 states that research should be interpreted in a broad manner, including fundamental and privately funded research. Additionally research in the public interest in the field of public health. Hence, these are two different origins. Research as such, which can be privately funded, and research in the public interest in the field of public health. It should be stressed that when the GDPR mentions public interest and research in the same sentence, it rarely does so as a combination. There is nearly always a comma between 'public interest' and 'research' indicating that thy are separate legal grounds.

[7] The concept of 'public interest' is used in discussions on politial philosophy and the legal debate. In both it can be used as a yardstick to measure what kind of decisions could be considered as being in the public interest or as the outcome of a procedure: if that procedure considers a certain decision or state of affairs in the public interest, it *is* in the public interest. The discussion then moves to what kind of procedure could lead to such an outcome. The latter conception is most of all used in the legal debate in the EU context where democratically chosen ofr democratically accountable public bodies may decide what is in the public interest. That conception should also be used to describe 'public interest' in the GDPR. But also then the outcome will be subjected to the rule of law, especially when invoking the public interest is used as an exeption to the general rule, as is the case here. The decision that a certain state of affairs / processing of personal data is in the public interest will be formally left tot the discretion of member states. Yet, how that is implemented should meet the criteria of necessity, proportionality and subsidiairity, which in extreme cases could claw back on whether the decision is in the public interest at all. See in the context of the TFEU and its predecessors the jurisprudence about the rule of reason and in the context of the ECHR the jurisprudence around article 8 of the ECHR.

			Privately funded is not defined. By charitable organisations can be considered privately funded but so could by pharmaceutical industry. That industry can be subsumed under research as meant in the GDPR also follows from the reference in the Recital to article 179.1 TFEU[8] I have read an interpretation of Recital 159 that research means publication of the data. This does not follow from the recital. It says that specific conditions should apply to the publication when that involves personal data[9]. This leaves a meagre picture. Though not explicitly stated in the GDPR but which in my opinion would follow from the debate around the special status of research in the GDPR, that research should (at least) fulfil the following partially overlapping conditions: • Research leads to generalizable results of the object of investigation; • Hence the data processing for research is as such is not meant to lead to decisions about specific data subjects[10]. There will always be a 'translation' of these results into daily practice[11].

[8] Treaty on the functioning of the European Union.

[9] Which goes without saying. Almost all research should be published as statistical findings. Yet, in biomedical research especially one might also want to publish pictures of the morphology of the disease which could be identifiable.

[10] With the exception of incidental findings which when they meet certain criteria should be reported back from a duty of care.

			• It is as much as possible reproducible[12]; • It will published in the public domain (which can mean scientific journals with payed access); • The underlying data will be FAIR[13] The last two dots are certainly more interpretative. But if one wants to use the research exemption, it should be research which adheres to present standards of transparent research. A 'grace period' for the application of patents can apply. To FAIR and connected datasharing the GDPR would again apply. That research should be ethical and vetted by an ethical review board (IRB) is not part of the definition. This definition is about what the claim for research in the context of the GDPR and research exemptions should mean. Not when it may be performed. Though that will come back when the research exemption is employed, either in general by national law (which would stipulate that a project is vetted first) or in each particular case
Personal data	Article 4.1		requires a longer discussion. Basically it means that non-personal data, hence anonymous data, cannot be easily be assumed. During the discussions about the GDPR the article

[11] It is admitted that especially in personalised or precision medicine (PM) research and clinical care tend to merge. But also in PM there is never a direct application of research findings (about possible valuable patterns based on more similar patients) to the patient but deliberation and discussion amongst researchers and physicians, sometimes further research and the interaction in the context of shared decision making with the patient who will ultimate decide about the possible implications of research. Another field where research and clinical care are often closely linked is rare genetic disorders. Yet, also here there is first the stage of finding a pattern even if they may be only one or two similarly affected patients in the world. Hence the importance of datasharing first, see also (e.g.): https://horizon-magazine.eu/article/sharing-data-between-researchers-too-often-afterthought-rare-disease-work-prof-hanns-lochm

[12] Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? Sci Transl Med. 2016 Jun 1;8(341):341ps12

[13] Findable, accessible, interoperable and reusable. See amongst others: https://www.dtls.nl/fair-data/

			29 Working Party issued an opinion on anonymisation techniques (Opinion 5/2014) which should be nuanced in the context of the final text of the GDPR and the decision on the ECJ in the Breyer case[14]. The case was about whether dynamic or floating IP addresses should be considered personal data in the German system but both the Advocate General and the Court issued much broader statements about whether data should be considered anonymous or not..
Pseudonymisation	Article 4.5		Again, requires a longer discussion. It should be mentioned that as also follows from the Recitals that the GDPR uses the phrase 'pseudonymisation' in a context where the controller would create a pseudonym. In that context they should still be considered personal data. The GDPR does not deal with situation where a third party (Trusted Third Party) would create a secure one way (one way hash, no possibility to go back to the data on which the pseudonym is based) pseudonym. Those data could still be considered anonymous data, if the data under the pseudonym are not indirectly identifiable, both when leaving the data source or when combined in the research domain. Which will often not be the case for more nuanced research data.
Purpose limitation	Article 5. b	Further processing for scientific research or historical research purposes is not incompati-	This clause which was also in the precedent Directive 95/46/EC, is of utmost importance. Especially in health research there is chain of data. Say from health care providers to a research database. The data were collected at

[14] Breyer, C-582/14, ECLI:EU:C:2016:779

		ble with the original purposes for which the data were collected ….	the source to deliver health care. Further processing to make those data suited for the research purpose, is by definition not incompatible with this original purpose. Three remarks here. • This clause is not dependent on implementation in national law. National law should on the contrary not contradict the essence of this clause. • If the personal data were to transferred to a third party, that party would of course need a separate legal ground to process those data. • Article 6.4 also holds a clause about further processing and compatibility, leading to a more nuanced (and restrictive) regime. It is submitted here in the case of research one does need to not fall back on 6.4 as that clause is meant for situations which are not covered by article 5 already. The more precise clause supersedes a general clause.
Storage limitation	Article 5.1.e Personal data shouldn't be kept in a form which permits the identification of subjects not longer	Exception if this necessary for scientific research purposes and in accordance with art. 89.1 and subject to appropriate technical	89.1 also refers to organisational and technical safeguards and adds more. We come back to 89.1 later in the scheme. One may wonder by the way whether this is really an exception as the basic rule already refers to 'necessary'. In combina-

	than is necessary for the purposes of processing[15]	and organisational measures	tion with the purpose limitation, it makes more sense. The original purpose would allow for anonymization. Yet, the added research purpose does not. In that case one does not need to fully anonymize the data, yet one does need to meet the additional safeguards.
Informed consent	Article 4.11; 7; • Consent should be specific • Consent cannot be implied in matters for which processing of personal data is not necessary; • Consent must be demonstrable by an affirmative action; • consent can be withdrawn.	Recital 33 It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain	Informed consent is at the basis of most health research. In that sense the second dots of the second row to the left were always met in research. Further processing for data collected in the context of a clinical trial is acknowledged in Regulation 536/2014 but needs a separate consent. The discussion was whether broad consent which especially has become the norm in biobanking, would still be feasible under the GDPR. Recital 33 accommodates these worries to a certain extent. It seems that participants can give broad consent to certain areas of research (given ethical oversight of that research) which in practice are sometimes very broadly defined (such as healthy aging[16]) but should be given the opportunity to narrow this down. The Recital certainly does not mean the opposite, meaning that consent should be narrow and (very) specific but participants should be allowed to broaden it. Such an interpretation

[15] It should be mentioned that this is one of those rather strange or at least clumsy formulations in the GDPR. If personal data are kept in a way which does not no longer permits the identification, they are not personal data anymore. But the intention of the clause is clear.

[16] UK Biobank, Lifelines, etc. all use broad consent and have been ethically vetted before being able to use broad consent.

		areas of research or parts of research projects to the extent allowed by the intended purpose.	would be contrary to Recital 33. Article 28.2 of the Clinical Trial Regulation (EU 536/2014) allows to consent for 'further use' of data collected in the context of trial outside the protocol for scientific research. That seems to imply broad consent. Yet, the article mentions 'without prejudice to Directive 95/46/EC . That Directive will be replaced by the GDPR under discussion here. 28.2 also mentions 'in conformity with applicable law on data protection' in the last sentence. Hence, the interpretation of broad consent in the GDPR also applies to the seemingly broader possibility to consent in the Clinical Trial Regulation.
Public interest in the field of public health	Article 9.2.i	Recital 157 stresses the importance of registries for research …"by coupling information from registries, researchers can obtain new knowledge of great value…" ….. "can provide the basis for knowledge-	Provides an exemption to the informed consent principle. Must be based on national or Union law[17]. Legislation must provide safeguards for the rights and freedoms of the data subject, in particular professional secrecy. There are many examples of the application of this clause. All EU countries should have implemented the WHO regulations on communicable diseases. The Scandinavian countries have disease registries. Some countries have those as well on the national or regional level, such as cancer registries. A publication in the Eur. J. Cancer shows how full

[17] It should be mentioned that processing for the health care or social system is also exempt from the informed consent requirement of the GDPR when be based on national law (where informed consent can again play a role obviously) and when the data are processed under the responsibility of a professional subject to the obligation of professional secrecy.

		based policy, improve the quality of life for a number of people…".	attempts for full anonymization of these registries as was attempted in Germany will lead to false results[18]. Recital 157 stresses the importance of these registries beyond statistics of incidence and prevalence for health research.
Processing necessary for scientific or historical research purposes	Article 9.2.j	Recital 156	9.2.i allows for exemption of the informed consent principle for research. Recital 156 also encompasses various other provisions of the GDPR as will be discussed later. Must be based on national or Union law. 9.2.h refers to article 89.1 which will be discussed further in the text. The exemption to informed consent in national of Union must be proportionate, respect the essence of the right to data protection and provide for suitable safeguards. See hereinafter at article 89.1
Further conditions/limitations	Article 9.4	Member states may maintain or introduce further conditions/limitations with regard to	Meaning as yet unclear, also considering that member states can regulate that the prohibition to use such data cannot be lifted by informed consent (9.2.a) such as is the case in many countries regarding testing for certain private life insurance.

[18] Andersen MR. et.al. 'Cancer registration, public health and the reform of the European data protection framework: Abandoning or improving European public health research?'. European Journal of Cancer (2015); 51. 1028-1038.

		processing genetic data, biometric data and data concerning health	An example might be the following. We have understood that in France a data processor for health data should be approved by the CNIL *and be established in France*. It remains to be seen whether the latter kind of restriction remains with the boundaries of article 1.3 about the free flwo of information in the EU.
Processing which does not require identification	Article 11	In the chain of data[19] for research identities of the participants are usually masked. That's privacy by design. Does not always mean that data will be anonymous and hence subjects rights would still apply. Not always, see the text on the right.	Article does not *specifically* relate to research but is mentioned here as the importance of this article is often overlooked. The article states that if the purposes of the processing do not require the identification of the data subject by the data controller, (and the controller does not have the kind of identifiable data by which the data subject can be contacted), the controller does not have to re-identify the data subject to comply with the GDPR (such as notification that data are being processed). This article solves the paradox that in order to comply with the GDPR one should know the (direct) identifiers of the data subject and hence retrieve those while one does not want to know them and the privacy by design is arranged in such a way that one cannot reasonably know them.

[19] There usually is a chain, from source data with often intermediary zones to the first research database and other research databases after FAIR to outcomes of analyses. Kuchinke W. et.al. 'A standardised graphic method for describing data privacy frameworks in primary care research using a flexible zone model'. Journal of Medical Informatics (2014) 83 941-957.

			If the data subject would provide additional information by which the controller would be able to retrieve the data subject in the database, rights of the subject as in 15-20 GDPR will again be applicable (unless a research exemption would apply).
Transparency when data have not been obtained from the data subject	Article 14 Data come from another controller, new controller must notify the data subject etc.	14.5.b If provision of such information would be a disproportionate effort, such as for scientific research but then subject to the safeguards of 89.1 *or* in so far as disclosure would seriously impair or make impossible the objectives of the processing.	A balancing act here. Exemption from disclosure specifically for research is subject to article 89.1 and hence see later. The impossibility or disproportionate effort is not. Research is mentioned as a particular example and then further conditions apply. In all cases appropriate measures must be taken to protect etc , *including making the information publicly available* In principle, data subject could be aware that data can be transmitted for research by the notification of the first controller, see 13. 1.e
Right to erasure (right to be forgotten)	Article 17	Does not apply … For reasons of public interest in the area of public health pursuant to 9.2. h and i (17.3.c)	Article 89.1 again, see later

		For …research in accordance with 89.1 and insofar as research would be rendered impossible or seriously impaired (17.3.d)	
Right to object	Article 21 Several grounds for the right to object	Also when data are processed pursuant to art. 89.1. But not if processing is necessary for a task carried out in the public interest (21.6)	This is the only example where research and public interest squarely merge. This would mean that – in the absence of EU or national legislation, see the discussion at 89.2) - there should always be a possibility to object to data processing in the context of 9.2j and 89 (hence without consent) *unless* the research is performed for a task in the public interest. The latter would mean that public authorities (as defined by member state law): • have assigned a task to an entity, • have explicitly considered this task in the public interest[20] an example would be monitoring and further research (there would always be a blurry line between the two[21]) of

[20] The concept of 'public interest' has various meanings depending on the context of the debate. In that of politics it is (rhetorically) forwarded as a standard to achieve and is juxtaposed to private interests or organising society according to (neo)liberal principles. In political philosophy it also a standard yet which gets more substance given a specific political theory. For a discussion see V. Held, The public interest and individual interests, Basic Books, Ney York/London, 1970 . What matters here is a legalised conception of the public interest as an exception to a general rule. In that context (of EU law, see the rule of reason doctrine, but also in that of the ECHR) it is always a democratically accountable public authority (which is always deemed to act in the public interest in general) that can specifically invoke the public interest (sometimes in retrospect when challenged but in the case of this article in advance) as a legitimation of this specific state of affairs. It will have a margin of discretion then within general more objectively defined boundaries of necessity and proportionality.

			antibiotics resistance of patients.
The general exemption	Article 89.1	Processing for research shall be subject to appropriate safeguards, ..for the rights and freedoms of data subject… data minimisation …pseudonymisation provided that purposes can be fulfilled in that matter. .. when by further processing purposes can be fulfilled by data which do not longer permit identification of data subjects, it should be done in that manner	There is something paradoxical in the clause. As these criteria should apply to <u>all</u> data processing whether for research or not. Read without context it might seem as if extra stringent criteria apply to research. Given, amongst other things, that research is also mentioned in the EU Charter (article 13) as one the freedoms and also considering article 179 TFEU, that seems an unacceptable conclusion. In the context of the research exemptions discussed above, especially 9.2.j, the clause gets a more substantive meaning. Necessity, proportionality and subsidiarity of processing of personal data for research must be accounted for, *especially* if not based on informed consent. The research exemption is not a 'carte blanche'. Neither would informed consent by the way. But without informed consent the research protocol should defence the data chain[22] even more. Especially why in the early stages of the chain informed consent is not a feasible option. Though that criterion is not mentioned in 89.1, it should be seen as part of the 'proportionate' mentioned in 9.2.i. The other criteria mentioned in 9.2.j are more difficult to

[21] For example by delving deeper into the characteristics of the bacteria, the previous antibiotic treatment or contacts of the patients with other patients who were possibly exposed to multiresistant bacteria. It might be argued that without such additional data the monitoring registry has little added value for health protection.

[22] See footnote 18

			phrase in more exact terms. Transparency can be considered one of them (see also the last sentence of 14.4.b) and the right to object unless the research would also be necessary for tasks carried out for reasons of public interest (see 21.6) or there is more general exemption to the right to object in the context of research laid down in EU or national law.
Exemption specific articles	Article 89.2	Union of member state law may provide derogations of 15, 16, 18, 21 for research	Research must then meet the conditions of 89.1 and the exemption must – to summarize – meet the criterion of necessity (likely to render impossible or would seriously impair the research) These exemptions come in addition to the exemptions discussed above. With the exception of the informed consent exemption, those discussed above are not dependent on national or Union law but provide nuances within the system of the GDPR. Those in 89.2 must be implemented in EU or national law and relate to: • right of access (15); • right to rectification (16); • right to restriction of processing (18); • right to object (21)[23].

[23] Trying to clarify the relation a bit further. If there is informed conset than there can also be withdrawal of consent and one does not neet to fall back on the right to object. Hence it seems that the right to object only makes sense in the context of national or EU law which grants a research exemption to informed consent. If that exemption would not also exclude the right to object, the directly applicable right to object (21.6) would remain in force, unless the research is necessary for a task carried out in the public interest. The latter exemption does not need to be laid down in national law as it is also directly applicable (yet, that this is a task carried out for the public interest, must follow from national law). If national or EU law would also

			In the latter case it seems that national law may derogate to the restriction of 21.6 discussed above.

exclude the right to object, then there is no right to object at all, even if the research is not carried out for a task in the public interest. Obviously in both cases the research exemption must fulfill the criteria of necessity, proportionality and subsidiairity.

This is version 1.4 of 16 November 2017. It might be superseded by later versions. Please check the website.

The text of version 1.4 is composed by Evert-Ben van Veen, Ll M. Evert-Ben van Veen is senior consultant at MedLawconsult and director of the MLC Foundation. MedLawconsult gives hands on legal advice in health care, most of all public health. The MLC Foundation supports health research with data and the discussions around a learning health care system.

Compared to version 1.3 in version 1.4 the MedLawconsult logo was added and the funding for the research leading to this overview.

Contact: eb.vanveen@medlaw.nl