GDPR facts and comments: LIBE and research
Comparison between the Commission proposal and that of LIBE

Evert-Ben van Veen L.I.M.[1]

Summary

LIBE:
1. broadens the definition of personal data.
2. brings pseudonymised data within the scope of personal data and hence the GDPR, but conflates two concepts there:
   a. The original meaning of pseudonymised data where pseudonymisation is a privacy enhancing technology to mask the identity of the subject by the data source for the recipient of the data;
   b. Where the recipient receives the data directly from the data subject by an electronic address and can single out and reach the data subject through that address.
3. Tightens the conditions for informed consent.
4. Makes art. 10 (stating a for research very important exception to rights of the data subject) internally self contradictory.
5. Adds many more regulatory provisions to the processing of personal data.
6. Broadens the powers of the coming European Data Protection Board without assuring proper governance of that EDPB.
7. Makes the conditions for data exchange with non-EU countries more complex and, together with the broadened definition of personal data, would severely hamper research collaboration with countries such as the U.S.
8. Almost completely abolishes the exemption for research with health data, as defined by LIBE, without informed consent, as defined by LIBE. Hence LIBE would be the end of registry research and other health research projects as exist in many European countries.

In the comments I explain and evaluate the LIBE proposals. In the final comments I state, amongst other things:
9. Patient data have a collective function next to an individual function, being to diagnose and treat the patient. It is through that collective function that the data could become available for the patient in the first place. They should be used to improve the further collective function, if certain strict conditions are met.
10. The quest of anonymisation before data reach the research domain is a vain one or leads research into a data desert. Data must be sufficiently detailed to find the delicate patterns which might predict what can keep us healthy and what not. Instead of on anonymisation techniques *before* data reach the research domain, the emphasis should be on how safety and confidentiality of data *within* the research domain can be assured. Medical scientific researchers should not be seen as aliens and potential adversaries of patient confidentiality, but as how they see themselves: part of the collective endeavour of health care to protect health and improve health care.

[1] This fact sheet has been made possible by a substantial contribution from COREON, the committee on regulatory affairs of observational research of the Dutch Federation of Biomedical Scientific Societies.

1. Introduction

Concerns have been raised about the report of the Committee on Civil Liberties (LIBE) of the European Parliament regarding the draft of the General Data Protection Regulation (GDPR). In the following, the articles of the GDPR[2] will be summarised as they were proposed by the European Commission (hereinafter: the Commission) and should be amended according to LIBE.
The proposal and the LIBE report are both very long documents and hence, if this document is to be easily accessible for researchers, it can only be a very brief summary. In the time between the Commission and LIBE there have been comments from the member states and other committees of the EP. These will only be referred to when strictly necessary to illustrate a certain point. I will follow the articles of the proposed GDPR. However, the GDPR has to be understood in context. That context is explained in the Recitals. Therefore, when necessary I will also quote from the Recitals. Another aspect of that context is the "constitutional" background of the proposals. I will give an extremely brief outline of that background in the following section.

I will try to be neutral in summarising the articles and give a more evaluative point of view in the comments at the end of each section. A caveat: the articles discussed are normative texts. Every summary has an aspect of interpretation and cannot be purely 'factual'.

So, this factsheet has it's limitations but I do hope that it will provide a more concrete overview of what is coming, or not, if the scientific community together with patient organisations can prevent it. How to prevent it relates to the best strategy and to the political, practical and ethical arguments to be used in that discourse. This fact sheet is not meant to discuss that strategy but some of the possible arguments will appear in the comments.

As in this digital age the reach of this factsheet cannot be predicted in advance, there are also general comments about how research with data is being performed which researchers know already but could be helpful for those, such as MEPs, who probably don't. In some comments I could not refrain from making rather poignant remarks about the prevailing attitude of many DPA's. They might find that I underestimate their responsiveness to necessities of medical research for the public good. It is up to them to show that I am wrong.

[2] See the last page for a list of abbreviations

2. The 'constitutional' background,

Formally the EU does not have a Constitution. The working of the EU is based on treaties between the constituting member states. The EU can only operate within the authority delegated by the member states through those treaties. The Treaty on the Functioning of the European Union (TFEU) is most relevant here. The proposed GDPR is based on articles 16.2 and 11.1 of the TFEU.
Art. 16.2 relates to data protection by those institutions and/or regarding those activities which fall within the scope of EU law. Art. 16.1 states a general right to protection of personal data. Art. 114 relates to the approximation of laws etc. between the member states and the procedures to be followed then.
The EU has 'subsidiarity' enshrined in the Treaty on the European Union (art. 5.3) and Protocol 2 to that Treaty. In the GDPR debate 'subsidiarity' does not seem to play an important role (any more).

Most of what concerns us, is based on 16.1 TFEU (though not given as the basis, that article is mentioned in the first recital) and 114 TFEU, *in combination with* the EU Charter of fundamental rights (hereinafter: the Charter).
The relation between the Charter and EU regulations is a complex one. In short, EU decisions cannot be based on the Charter but they have to comply with it and can in principle be challenged in the European Court of Justice if they – allegedly – don't.

The Charter has a separate article on protection of personal data, article 8. It states the following:

1. Everyone has the right to the protection of personal data concerning him or her.
2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
3. Compliance with these rules shall be subject to control by an independent authority.

As with all fundamental rights instruments, these provisions state principles as well as hard rules and have at times to be balanced against other principles. Take article 13 of the Charter on the freedom of the arts and science. That states:

> The arts and scientific research shall be free of constraint. Academic freedom shall be respected.

The intention is clear but in such bold terms it is obviously nonsense when understood as a rule. This freedom has to be weighed against other principles.

The Charter also has an article on health care (art. 35). It should be noted that this article is not in the chapter on freedoms (as art. 8 and 13) but in that on solidarity. It does not state a 'right to health' as such, but states:

> Everyone has the right of access to preventive health care and the right to benefit from medical treatment under the conditions established by national laws and practices. A high level of human health protection shall be ensured in the definition and implementation of all Union policies and activities.

The article on data protection is detailed, seems to be much less stating principles which have to be balanced against other principles and is as such unique in international documents. It is the main source of inspiration for LIBE and of many others in the debate.

With this in mind, on to the changes to the draft GDPR as proposed in the LIBE report.

3. Definitions

Personal data

Two definitions are in combination most important here: 'data subject' and 'personal data'. That defines most of all the scope of the Regulation.

According to LIBE the definition of data subject should be (here and hereinafter, unless stated otherwise, in bold the changes to the Commission proposal):

> 'data subject' means an identified natural person or a natural person who can be identified ***or singled out***, directly or indirectly, ***alone or in combination with associated data,*** by means reasonably likely to be used by the controller or by any other natural or legal person, in particular by reference to a ***unique identifier,*** location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural, social ***or gender*** identity ***or sexual orientation*** of that person;

Personal data then means any data relating to a data subject

LIBE also proposes a change in the recitals, as follows:

> (23) The principles of protection should apply to any information concerning an identified or identifiable person. To determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the individual. ***This Regulation should not apply to anonymous data, meaning any data that can not be related, directly or indirectly, alone or in combination with associated data, to a natural person or where establishing such a relation would require a disproportionate amount of time, expense, and effort, taking into account the state of the art in technology at the time of the processing and the possibilities for development during the period for which the data will be processed.***

Comment

With respect to the system of Directive 95/46/EC the wording of the 2 concepts has been reversed. Personal data has a long definition in the Directive, also in relation to recital 26 of the Directive; data subject a short one.

LIBE argues that it's definition is consistent with earlier Opinions of the art. 29 Working Party[3] (especially 4/2007).

That is not the case. The concept of what constitutes personal data has been (considerably) broadened

Opinion 4/2007 makes clear that also 2 way pseudomised data *can*[4] be anonymous data. As, though admittedly not altogether clear on that, the WP acknowledges that there can be a data chain (and there is always a data chain in research) where data which are to be

[3] The Working Party is the assembly of the national DPA's as constituted by art. 29 of Directive 95/46/EC. The WP issues 'Opinions' and some of those will be referred to in the text.

[4] There are two aspects which should be taken into account then:

- safety of the pseudonym (no possibility for 'replay back');
- the level of detail of the data under that pseudonym (not indirectly identifiable)

It should be mentioned that already under 95/46/EC the threshold for both tests is extremely high. Hence in the original (see the text) conception of pseudonymised data there can be:

- pseudomyised – anonymous data, *and*
- pseudonymised non anonymous data.

For a lengthy discussion on these two aspects see E. B. van Veen, Patient data for Health Research, October 2011, available via: http://www.medlaw.nl/?p=435

considered identifiable at the source can be anonymous data at the recipient because of the 2 way pseudonymisation.[5]
Under the GDPR the Commission considers that type of data personal data. The Commission stresses 'or by any other person' in that context. However, 'or by any other person' could also be understood as a person who has access to the data at the recipient. And in the context of research those data will often be different from the data from each of the sources as data are collected in the research domain under a pseudonym from different sources.

In most comments it was said that the definition was too broad and that pseudonymised data should be exempted from the GDPR or at least have a 'status aparte' with a softer regime.

LIBE however, does not give a different status to pseudonymised data (at least not insofar as relevant to research) and broadens the definition even more.
It also adds a definition of pseudomised data, as a subset of personal data, being:

> ***(2a) 'pseudonym' means a unique identifier which is specific to one given context and which does not permit the direct identification of a natural person, but allows the singling out of a data subject;***

All subjects in research databases need to be 'singled out' in the sense of being discerned from each other. If that is meant, there are hardly any anonymous data left in the context of research.
LIBE uses pseudonymised data also in the context of the articles on research (81.2 and 83, see hereinafter) hence, this broad definition seems to be meant.

However, "singling out" could also have a more limited meaning, being 'singling out', as is said in Recital 21: *"particularly in order to take decisions concerning her or him or for analysing or predicting her or his personal preferences, behaviours and attitudes."* This was also the background of the art. 29 WP in its latest Opinion on the draft GDPR. The WP proposed mentioning 'singling out' as an element of personal data, *however*, in the context of making decisions regarding that person who is 'singled out' or "evaluating" him or her (WP 8/2012, p. 6). This evaluation seems to refer to 'personalised advertising'.
Decisions or personalised advertising without knowing who that person is, can for example be made/done based on an IP address or the IMEI of a mobile telephone. And not only can the person (or persons: the pc might be shared etc.; for research such singling out would in principle be meaningless) behind that digital address be 'singled out', he or she can also be reached through that address.

The data in a research database, whether pseudonymised or not, are not used in such a way, but to find patterns, which can explain general explanatory correlations between a combination of – in short – environment, life style, disease, often – omic data, treatment and health. It is about "patterns, not persons". In the research domain one person is as such never 'evaluated' but always together with many, even thousends of other persons according to predefined search questions to find those patterns. Yet, such patterns can never be found in a 'datastew' where persons and their attributes cannot be discerned from each other. The results are general, statistical data and will be used in health care in

[5] Also ISO norm 25237 (2008) allows for two-way pseudonymisation leading to anonymous data.

general, for persons who share the identified characteristics. For the subjects in the database such explanation might even come too late.[6]

Hence LIBE's definition of 'pseudonymised data' conflates two concepts:

- Singling out in the sense of the last Opinion of the WP, namely to 'evaluate' a person behind an electronic address based on the data transmitted by that person through that address and with the possibility to reach that person through that address with content based on that evaluation[7]
- A privacy enhancing technique (PET) in a data chain to distinguish persons from each other by the end recipient and to match data from different sources to the same person, yet, without knowing who they are and not being able to reach them, in order to find through the aggregate of the data of all those persons meaningful correlations which will result in general statistical results.

The conflation of the two concepts creates difficulties later.
The first conception of pseudonymised data is indeed a new branch of personal data with possible privacy infringements attached and could need special regulatory attention. The holder of the pseudonymised data has received the pseudonym (electronic address) directly from the subject concerned and can in principle reach that subject through that pseudonym.

The second conception is a PET which *enhances* privacy. The pseudonymisation is applied in a data chain. The source or a Trusted Third Party replaces the direct identifiers of the subject by a pseudonym. The recipient cannot reach the subject through that pseudonym. The data attached to the pseudonym will in general be filtered to prevent all too easy re-identification through those data. In the meaning of ISO 25237 (nt. 5) of pseudonymised data, they should even be anonymous then. This second conception is the *original* meaning of pseudonymised data as it is, amongst others, used in the context of research.

Consent
What constitutes consent has been made much stricter:

> 'the data subject's consent' means any freely given specific, informed and explicit indication of his or her wishes by which the data subject, either by a statement or by a clear affirmative action, signifies agreement to personal data relating to them being processed ***for one or more specific purposes***;

Comment
Explicit consent for processing health data was already – outside the exemptions - necessary under 95/46/EC. The threshold for valid consent has been made much stricter and even more so by LIBE. This does not help for broad consent for research which has become more and more the standard when consent is needed.

Additionally it should be mentioned that LIBE adds the following to the Recitals:

> Where processing is based on the data subject's consent, the controller should have the burden of proving that the data subject has given the consent to the processing operation. In

[6] In tissue bank research, results of 2 way coded data could also be used for individual feed-back to the subjects concerned. Only in that case, there would also be 'singling out' in the strict sense of the word .
[7] The WP is not altogether clear on this. The first is needed for the second but the second does not necessarily follow from the first. If the first would already amount to processing of personal data with all the strings attached to that, especially according to LIBE, it might become difficult for service providers to adjust their services to the statistics of client behaviour, such as that rush hours and allocating bandwidth to that aim in a business district are different from residential areas. In the first phase of those statistics the electronic addresses cannot be excluded.

> particular in the context of a written declaration on another matter, safeguards should ensure that the data subject is aware that and to what extent consent is given. ***To comply with the principle of data minimisation, the burden of proof should not be understood as requiring the positive identification of data subjects unless necessary.***

And:

> In order to ensure free consent, it should be clarified that consent does not provide a valid legal ground where the individual has no genuine and free choice and is subsequently not able to refuse or withdraw consent without detriment. ***The use of default options which the data subject is required to modify to object to the processing, such as pre-ticked boxes, does not express free consent.***

Opt-out systems are hence 'ousted'. Such op-out systems exist for example in many European countries, based on statutory provisions, to use health care data for research or regarding using residual tissue and attached data for research, such as, in the latter case, in Denmark and Belgium.
While in certain countries the system was meant to use only 'pseudonymised –anonymous data' for such research, the new definition of personal data by LIBE brings such data into the consent system as well. I will come back to that when discussing art. 83.

4. General principles

Article 5 states general principles of data processing: transparency, purpose limitation, data minimisation (not more than necessary and proportional to the aim), data integrity, accountability. 5.e also makes a certain exemption to the storage period:

> Kept in a form which permits identification ***or singling out*** of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the data will be processed solely for historical, statistical or scientific research purposes in accordance with the rules and conditions of Article 83 and if a periodic review is carried out to assess the necessity to continue the storage ***(storage minimisation)***;

Comment
This change seems the least problematic to me, apart from the problems of art. 83, see hereinafter. Of course you should have a rationale to retain the data. And there are many in the context of research: from validating older research to building upon old data for new research (without harassing data subjects again) as in the context of datasharing.

A new article is proposed by LIBE being:

> *(5.)****1a. Processing of personal data shall be organised and carried out in a way that ensures compliance with the principles referred to in paragraph 1; producers, data controllers and data processors shall take technical and operational measures to ensure such compliance in the design, set-up, and operation of automatic data processing or filing systems.***

This is an application of the so called “privacy by design” principle. Privacy by design comes back on many occasions in the GDPR and is strengthened by LIBE. It must also shown to be done, by record keeping etc. I will not repeat that in the following.

Article 7, being the conditions for consent, did not change, insofar as relevant for research. Those conditions are quite elevated already. See the above.

Article 9
This article relates to special categories of personal data. LIBE broadens their scope:

> The processing of personal data, revealing race or ethnic origin, political opinions, religion or ***philosophical*** beliefs, ***sexual orientation or gender identity,*** trade-union membership ***and activities***, and the processing of genetic data or data concerning health or sex life or criminal convictions***,*** or related security measures shall be prohibited.

Comment
LIBE does not substantially change the conditions for exemptions to the prohibition on processing those special data. Example given, 8.c (processing to protect the vital interests of the data subject or someone else who cannot give consent) is extremely important in cases of possible child abuse and public mental health. It remains intact. At certain articles LIBE adds ‘adequate safeguards for the fundamental rights of the data subject’.

Article 9 also refers to the exemptions on the consent principle in art. 81 and 83. LIBE does not change art. 9 in that respect but changes art. 81 and 83 substantially, as will be explained later.

Article 9, paragraph 3

LIBE amends this as follows:

> ***The European Data Protection Board*** shall be ***entrusted with the task*** of further specifying the criteria, conditions and appropriate safeguards for the processing of special categories of personal data referred to in paragraph 1 and the exemptions laid down in paragraph 2***, in accordance with Article 66***.

Comment

The EDPB will be the successor to the art. 29 Working Party (see footnote 3). It will get a status as a European Institution under the GDPR. LIBE enters the EDPB on many more occasions. Usually that the Commission should consult the EDPB before adopting an implementing or delegated act. Here, and at certain other articles, the EDPB gets even more authority. Not the Commission should issue further rules but the EDPB. This is potentially worrying. Many comments have been made on the possibility for 'delegated' and 'implementing' acts which the Commission attributed to itself on manifold occasions in the GDPR. All earlier comments stated that most of these possibilities should be repealed. LIBE does much less so and sometimes, as in the present case, substitutes the EDPB for the Commission.

With by-laws by the Commission there is some system of oversight, as regulated in the TFEU and following decisions. For implementing acts the Commission must consult an expert group. Delegated acts can even be repealed by the EP or the Council (member states).
There is non of that governance system with the EDPB. In its present form, the art. 29 WP, the Opinions are not discussed beforehand with relevant stakeholders. EDPB will be composed of the DPA's of the member states, just like the present art. 29 WP. The EDPS will be added. Recruitment for and socialisation in those functions usually means that they have a rather one sided view on the issues at stake. There is nothing in the articles on the EDPB that states that stakeholders, other than the Commission, should be consulted before the EDPB makes a decision. Contrary to decisions of the national DPA's there is no redress system either.
Governance of the EDPB is a serious issue.

Article 10

LIBE amends this as follows:

> If the data processed by a controller do not permit the controller to identify ***or single out*** a natural person, ***or consist only of data relating to pseudonyms,*** the controller shall not be obliged to acquire additional information in order to identify the data subject for the sole purpose of complying with any provision of this Regulation

And gives the following justification:

> *Data controllers may use a unique identifier for the same person across different services and contexts, while still not being able to identify a natural person on their basis. Pseudonyms as defined in the amendment to Article 4 are limited to a specific context. The amendment makes clear that the article applies to both cases*.

Comment

In the EUROCOURSE comment[8] to the draft GDPR it was discussed how important this article is for research. The right to be informed about processing if the data are not

[8] Available at: http://www.medlaw.nl/?p=511

assembled from the data subject directly, to have a copy of the data, that of erasure, etc. (see the next paragraph) do not apply when the controller can fall back on this art. 10, hence cannot reasonably identify the subject.
LIBE does not clarify but adds considerable confusion. Art. 10 has even become self-contradictory.
The scope is according to LIBE first of all data which:

- Do *not* permit to identify or single out.

But then LIBE adds:

- *Or* (my italics) are only data relating to pseudonyms

Yet pseudonymisation always means 'singling out' according to LIBE's definition.

5. Rights of the data subject

These rights are, in summary:

- To be informed about data processing in a detailed way, unless that information was given already at the time when the data were collected from the subject;
- To be informed upon request about what data are being processed;
- Data portability;
- Rectification;
- To be forgotten and personal data erased.

Many of those provisions have been discussed in the EUROCOURSE comment already. All research databases will be subject to them, unless they can rely on article 10.
LIBE makes those more elaborate. What is revealing in this context is how LIBE proposes to change recital 50 (and here the bold lines are what LIBE wants to have deleted):

> However, it is not necessary to impose this obligation where the data subject already disposes of this information, or where the recording or disclosure of the data is expressly laid down by law, or where the provision of information to the data subject proves impossible or would involve disproportionate efforts. ***The latter could be particularly the case where processing is for historical, statistical or scientific research purposes; in this regard, the number of data subjects, the age of the data, and any compensatory measures adopted may be taken into consideration.***

LIBE adds:

> The deleted text may be misunderstood as promoting a lower level of protection for certain kinds of data processing.

Yet LIBE does not change art. 15.5.b which states that the data subject does not have to be informed when the data are not collected from the subject and informing him/her would be impossible or would involve a disproportionate effort.

Comment
It should be noted that these rights would also apply if one can fall back on the exemption to informed consent of art. 83 but not on article 10 as well.
Though in my opinion transparency should always guide research, some of those rights can give complications. E.g., the right of erasure can make it impossible to validate earlier research findings. There is more, as discussed in the EUROCOURSE paper.

The proposed rights and regulatory strings will probably have many more effects outside research. See the comments at the next section.

6. General articles about processing of personal data

There is too much to mention. Privacy protection by design and by default remain intact of course. Documentation of all processing operations would also apply to small firms (art. 28.4. b deleted). The necessity for a Privacy Impact Assessment (PIA) is further clarified. Health research would fall under the PIA rule anyhow. According to LIBE the PIA must also clarify the necessity and proportionality of processing personal data. Article 33.4 on how the PIA should be performed is amended as follows (and here the bold lines mean what LIBE wants to have deleted in the text):

> The controller shall seek the views of data subjects or their representatives on the intended processing***, without prejudice to the protection of commercial or public interests or the security of the processing operations.***

Under certain conditions the controller must consult the DPA first (art. 34.2). The DPA shall establish a list of processing operations which are always subject to prior consultation (art. 34.4). LIBE changes the DPA here into the EDPB.

The role of the data protection officer (DPO in the jargon) is strengthened. The professional qualifications of the DPO are spelled out and a DPO should now be appointed by a legal person with data processing which relates to more than 500 data subjects per year, such as a general practitioner or a small consultancy firm like mine.

On the Codes of Conduct LIBE proposes to amend art. 38.4 as follows:

> The Commission ***shall be empowered to adopt, after requesting an opinion of the European Data Protection Board, delegated acts in accordance with Article 86*** for deciding that the codes of conduct and amendments or extensions to existing codes of conduct submitted to it pursuant to paragraph 3 ***are in line with this Regulation and*** have general validity within the Union. ***This delegated act shall confer enforceable rights on data subjects.***

Hence, a European research Code of Conduct which would only elaborate on art. 83 would probably not be approved.

Comments

LIBE will unleash a regulatory avalanche which in my opinion will suffocate innovation and SME's. Researchers might find this not their main concern. Yet the coming slow death of innovative businesses and SME's and the direct blow to research for the public good follow from for the same regulatory approach and paradoxically governments might be more interested in the former in their discussions with the EP and the Commission. For the research lobby, it could be interesting to connect to those objections.

LIBE will create jobs as well. At the 2013 conference in Brussels on 'Computers, Privacy & Data Protection'[9] many were delighted by the LIBE report. The crowd was an interesting mix of true privacy believers and those for whom the GDPR and certainly in the LIBE version would provide more status, regulatory powers or business ventures, such DPO functions or offering PIA services. [10]

But one might wonder whether those are the jobs Europe needs.

[9] www.cpdpconferences.org

[10] Admittedly, that would become my business model as well but I hope that I don't have to. A responsible approach is to prevent the avalanche, not to wait and come to rescue once it happened.

In the late 90'ies of last century a school kid started a site on the internet where consumers could compare prices and services of mobile phone providers. He had seen people struggling to find the best offers and thought that this could be done smarter. He developed a sustainable business model around it at an age when his parents had to file his business at the Chamber of Commerce as he was too young. Such a business model has since then been copied by many others. In his late 20'ies now, Ben Woldring runs several successful internet firms. Starting in the attic of his parent's house and with pocket money instead of in a glass palace and with foreign investors, his business was not affected by the burst of the dot com bubble. A privacy complaint has never been reported.
Young Ben Woldring could never have started his business under LIBE but probably also not under the original Commission proposal.

7. Transfer to third countries

Third countries are non-EU countries. In principle, any transfer of personal data to third countries is forbidden unless one of the exceptions would apply.
Researchers collaborate with third countries, such as the USA. Under the Directive 95/46/EC it could be held that only anonymous data were exchanged. As seen, in the GDPR the definition of personal data might be considerably expanded. One usually does not collaborate with exchanging statistical information or a 'datastew' in which individuals cannot be discerned but often with datafiles in which they can, hence would be probably 'pseudonymised data' according the definition of LIBE.

Therefore the exceptions to the ban on transfer to third countries would become more important. These are, very much abbreviated:

- When the Commission has decided that the third country offers an adequate level of protection (art. 41);
- The Commission can also decide that a third country does not have such an adequate level of protection (41.5);
- When no decision has been made, transfer is possible, in the context of appropriate safeguards (42), being:
    - Binding corporate rules;
    - Standard protection clauses adopted by the Commission or a DPA;
    - Contractual clauses between the controller or processor and the recipient as approved by the DPA.

Art. 43 gives further details about 'binding corporate rules'.
Art. 44 gives additional, case by case exceptions on the ban of transfer to third countries, such as when the data subject has consented after having been informed about the risks of such transfer.

LIBE adds to this, again very briefly, the following:
In the context of art 42, appropriate safeguards:

> **Those appropriate safeguards shall, at least: (a) guarantee the observance of the principles of personal data processing as established in Article 5;(b) safeguard data subject rights as established in Chapter III and provide for effective redress mechanisms;(c) ensure the observance of the principles of privacy by design and by default as established in Article 23;(d) guarantee the existence of a data protection officer pursuant to Section 4 of Chapter IV.**

Obviously, this export of the European rules and the DPO, does not make cross-Atlantic cooperation with other research institutions easier. There is also no exception or a more lenient regime for pseudonymised data (as originally understood).

Much more could be said about exchange with third countries but these seem to me the main worries for research at the moment.

8. Institutional rules

I will be very brief on those. The role and independence of national DPA's is strengthened. The EDPB is established. A mechanism for consultancy between DPA's and consistent decisions on EU level is instituted.
The problems of governance of the EDPB have been alluded to already. Some national DPA's suffer from the same.

9. Article 81

Article 81 is about health data. The article builds upon the exemption for informed consent mentioned in art. 9.2 and then states:

a. the general conditions under which such data may be used;
b. for which purposes.

Ad a:

In the wording of LIBE:

> ***In accordance with the rules set out in*** this Regulation, ***in particular*** with point (h) of Article 9(2), processing of personal data concerning health must be on the basis of Union law or Member State law which shall provide for suitable and specific measures to safeguard the data subject's interests ***and fundamental rights***, and be necessary for *(see ad b, EBV):*

LIBE also adds to the general conditions:

> ***When the purposes referred to in points (a) to (c) of paragraph 1 can be achieved without the use of personal data, such data shall not be used for those purposes.***

As the rationale LIBE gives:

> *Clarification that the principle of minimisation of the processing of personal data also applies in case it is regulated by Member State law. Health data is extremely sensitive and deserves utmost protection.*

Ad b:

As proposed by the Commission the purposes are , in short (81.1):

a) for medical care and management of health services, subject to an obligation of professional secrecy;
b) for reasons of public health such as protection against serious cross border health threats or ensuring high standards of quality and safety, inter alia for medicinal products and medical devices;
c) other reasons of public interest, such as social protection, the quality and cost effectiveness of settling claims in the health insurance system.

In art. 81.2 it is stated that processing of personal data concerning health necessary for historical, statistical and scientific reasons is subject to the conditions of art. 83. Art. 81.2 subsumes under these reasons 'patient registries set up for improving diagnoses and differentiating between similar types of disease and preparing studies for therapies'.

LIBE does not alter 81.1 but changes art. 81.2 completely:

- Processing of personal data concerning health which is necessary for historical, statistical or scientific research purposes, ***shall be permitted only with the consent of the data subject, and shall be*** subject to the conditions and safeguards referred to in Article 83.
- ***Member States law may provide for exceptions to the requirement of consent for research, as referred to in paragraph 2, with regard to research that serves an exceptionally high public interests, if that research cannot possibly be carried out otherwise. The data in question shall be anonymised, or if that is not possible for the research purposes, pseudonymised under the highest technical standards, and all necessary measures shall be taken to prevent re-identification of the data subjects. Such processing shall be subject to prior authorisation of the DPA …….***

Art. 81.3 states that the Commission issue delegated acts pursuant art. 81 as to 'the public interest' and the safety measures taken. LIBE adds 'after consultation of the EDPB'.

LIBE further adds that member states must notify the Commission about its legislation pursuant to art. 81.1.

Comment
About 81.2 see further the next section ad art. 83.
Here only some short comments.

- The health care system of member states as such is outside the competence of the EU (art. 168.7 TFEU). Art. 81.1 should be read as a compromise between not 'trespassing' on the competence of member states regarding their health care system and yet issue norms about how data within that health care system should be handled;
- The EU is competent for setting high standards for certain products used in health care (art. 168.4 TFEU). Combined with the harmonisation of laws to bring about the free exchange of goods and services within the EU, the EU has issued norms about medicinal products (amounting to a very complex piece of legislation) and medical devices. In medical devices we see the trend from Directive (which must be implemented in the laws of the member states) towards a Regulation (which has direct binding effect) as well. Both state that PMS of such products must be performed. Hence their explicit mention in 81.1 b).
- The relation between 81.1 b) and 81.2 is intriguing. Many registries are first of all set up for quality assurance (providing bench marks with the aggregate of the data).[11] These would fall under 81.1 b). However, once the data are in the registry, they can be used for scientific research as well. That shouldn't change the status of those registries, if certain conditions are met on how the research on the aggregate of the data is performed.
- There are more provisions in EU legislation where member states must send their national legislation to the EU. Sometimes it even has the consequence that without such notification, that legislation is not valid. However, notification procedures always relate to 'implementing legislation'. Not to the type of legislation here which is about how member states have arranged their health care system and professional norms regarding medical confidentiality in that context. That the member states should send the full body of their (existing) health legislation to the Commission may without doubt provide additional work at both ends of the process but the added value is unclear to me.

[11] See slide 6 of my presentation on PMS of medical devices, available at: http://www.medlaw.nl/wp-content/uploads/ebvvtilburg2013-v2.pdf

10. Article 83

In the proposal of the Commission art. 83.1 gave room for research with personal data without consent if:

a) The research cannot otherwise be fulfilled (hence, with anonymous data);
b) Identifiable data are kept separate from the other data, **insofar research can still be performed in this way.**

LIBE proposes:

- To limit the scope of art. 83.1 to data which do not pertain to a child or are not sensitive data as explained earlier.
- to delete the bold lines of 83.1 b)
- to add (in line with LIBE's art. 81.2):
  - ***Subject to the exception in paragraph 1b, data falling within the categories of data covered by Articles 8 and 9 may be processed for historical, statistical or scientific research only with the consent of the data subjects.***
  - ***Member States law may provide for exceptions to the requirement of consent for research, as referred to in paragraph 1a, with regard to research that serves an exceptionally high public interests, if that research cannot possibly be carried out otherwise. The data in question shall be anonymised, or if that is not possible for the research purposes, pseudonymised under the highest technical standards, and all necessary measures shall be taken to prevent re-identification of the data subjects. Such processing shall be subject to prior authorisation of the DPA ……..***

Art. 81.3 states that the Commission can adopt delegated acts on the issues raised by art. 81.1 and 2. LIBE proposes to delete 81.3. This is logical as the EDPB will in LIBE's view be empowered to issue further decisions pursuant to art. 9 (see section 4). Article 83 builds on art. 9.

Comment

We have three problems combined here:

- The definition of personal data is broadened;
- The threshold for valid consent is heightened;
- The possibilities for research with data of a child or with sensitive data without consent are very much limited:
  1. They must 'at least' be 'pseudonymised under the highest technical standard';
  2. The research must serve an 'exceptionally high public interest';
  3. All necessary measures must be taken to prevent re-identification of the data subjects;
  4. The DPA must approve of the processing.

Ad 1

LIBE mentions 'anonymised or at least pseudonymised' . In its eagerness to issue strict norms LIBE seems to forget that anonymised data fall outside the scope of the GDPR anyhow. But that's not the main point.

It should be admitted that all processing should adhere to what I call the NEPROS formula:

- Necessary;
- Proportional to that necessity;

- After a test of subsidiarity: can we not achieve these aims in another way, with less personal data and/or with consent.

Pseudonymisation can be an important tool in the NEPROS formula but will not always be feasible, especially in the first stages when data have to be collected at the sources.

Ad 2

This is a very worrying formula which leaves room for unfettered discretion by DPA's who have never done empirical research and tread on in the same ideological vein for decades. In a free society, every research with data which aims to bring its results into the public domain, is by definition in the public interest. There shouldn't be a watchdog which decides which we are allowed to know about our present social and medical practices, about the human condition in general, and what not.[12] In the following discussion in the professional and lay press and possibly by the legislator can then be decided what we choose to do or not do with those results.

Ad 3

The question is what 'constitutes all necessary measures'. I will come back to that in my final comments.

Ad 4

It could be argued that some review of a research proposal is necessary whether NEPROS is adhered to. However, a research ethics committee is much better equipped for that task, where member states will be free to structure that system and to avoid that one proposal, e.g. using data from various sources, must be reviewed by several committees. Also other 'stapling' of reviews, such and an ethics committee and the DPA, must be avoided.

For the non sensitive data LIBE deletes 'unless the research cannot be carried out otherwise'. However, often research cannot be carried out otherwise. The original version of the art. 83.1 is sloppy in its wording as well. It refers to what researchers always try to do, namely making a distinction between the data by which communication with subjects is possible or by which subjects can be discerned and data on which the actual research is done (see also footnote 4). However, the latter type of data can be indirectly identifiable as well, given the high threshold before data can be considered anonymous and the need to have a rich dataset which makes subtle correlations possible and by which confounders can be excluded.

Yet, all this by itself does not give sufficient basis for using sensitive data for research without consent. I will try to sketch that basis in the concluding comments, after some introductory remarks.

[12] Which research will be funded is a different question of course.

11. Final comments

1. Often researchers complain about the diverging rules for research in Europe and argue that these rules should be 'harmonised'. In response I have argued that this is a naïve view as usually such harmonisation leads to a levelling up to the norms of the stricter countries, and adds more (European) bureaucracy and barriers to research in countries with more lenient regimes.[13] Which does not mean that I am happy now that I might be proven right.
2. In my opinion, LIBE's proposals, and perhaps even the GDPR in general, often uses the fundamental rights parlance to 'protect us' against something which is not more than a - sometimes irritating and often very silly - nuisance, such as personalised advertising through computer algorithms based on our online behaviour. Big words like 'fundamental rights' should be used with caution. Otherwise we will have no words left for when it really matters.
3. The diligent use of patient data is indeed an area where fundamental rights are involved. Not only of the patients to whom those data relate but also of those who are dependent on advances in health care, as some patient organisations have made clear in their response to LIBE already.
4. As explained in my paper on patient data of health research[14], researchers are not by definition aliens to the intimate doctor-patient relation and the confidentiality which should reign there. They can be part of that context of privacy. Patient data have an individual and a collective function. The individual function is to use diagnostic findings as the patient seems best, according to his or her informed consent. Within the individual function we might grant patients granular control about how their data are going to be used. The collective function is what made these data available in the first place. A solidarity based health care system by which the patient could visit a doctor and receive care based upon over 2000 years of medical research and training which created the knowledge, expertise and techniques to diagnose and treat a patient. It may expected that patients who profit from that collective function, will contribute to it as well, for future patients. This is the underlying (ethical and legal) paradigm which explains why certain European countries have arranged research around data in their health care system with a lenient regime for such research.
5. Which certainly does not mean that everything goes. I have mentioned the NEPROS formula already. Explicit consent may be the golden standard but there are many occasions where consent is not feasible, would make research invalid or when it is even inappropriate to ask the patient for consent for 'further use' of data which relate to him or her. Under such conditions an opt–out possibility should in my opinion nevertheless be offered, in a general way, without bringing consent in again through the backdoor.
6. Of course, these data should be safe within the research domain. There is tendency in ICT land to devise anonymisation techniques as the panacea for using source health data, such as in electronic health care records, without consent in research. Much ICT research money is going in that direction. Often this ICT research is not done in close cooperation with medical researchers.
7. This approach is a dead-end street. Not the least as such ICT research is at the same time a challenge to others to show that certain datasets can nevertheless be re-identified under certain conditions. The race can never be won. The problem is not the

[13] See e.g. E.B. van Veen et al, 'TuBaFrost 3: Regulatory and ethical issues on the exchange of residual tissue for research across Europe', EJC 2006, 42 (17), p. 2914-2923, and P. Riegman, E.B. van Veen, Research with residual tissue, J. Human Genetics 2011, 130 (3), p. 357-368.
[14] Patient data for health research', The Hague, October 2011, available at: http://www.medlaw.nl/?p=439

safety of pseudonym but the data attached to each pseudonym. They must remain sufficiently detailed for health research, as explained earlier. And more and more data are available in the public or semi-private domain with which at least some of these research data could in theory be matched. In the paper mentioned I have proposed a different approach: instead of assuring that data reach the research domain anonymously (also if with a pseudonym attached), rather that they will not be re-identified within the research domain.

8. There are no publicly reported examples of such illegal re-identification of data by researchers but more can and should be done to assure data safety and confidentiality. ICT research should devise manageable and cost efficient database structures, also for smaller projects, where research with personal data can be performed under the conditions of sufficient safety against outside 'adversaries', as it is called in the jargon, and which will allow an audit trail of how researchers have handled their duty of confidentiality. Such structures should also allow for validating data and their provenance when the outcomes of research are challenged.
9. This approach does not see researchers as possible inside adversaries as in the LIBE report, but as how they see themselves: part of the common endeavour to protect health and deliver optimal health care. If the reactions to the LIBE report achieve that this understanding as it exists already in some European countries, will be become more wide spread, there will even be some unintended benefit from the LIBE report.

Abbreviations

COREON	Commission on Regulatory Affairs of the Dutch Federation of Biomedical Scientific Societies
DPA	Data Protection Authority
DPO	Data Protection Officer
EC	European Commission
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EP	European Parliament
EU	European Union
GDPR	General Data Protection Regulation
ICT	Information and Communication Technology
IMEI	International Mobile Equipment Identity
IP	Internet Protocol
ISO	International Organization for Standardization
LIBE	Committee on Civil Liberties, Justice and Home Affairs
MEP	Member of European Parliament
NEPROS	NEcessary, PROportional, reviewed for Subsidiarity
PIA	Privacy Impact Assessment
PMS	Post Marketing Surveillance
SME	Small and Medium Enterprise
TFEU	Treaty on the Functioning of the European Union
WP	Article 29 Working Party

This report can be shared by anyone who is interested in the coming GDPR and observational medical research in Europe.

Comments can be addressed to MedLawconsult, via the address provided on our website. As we are not a professional lobbying organisation and funds for this activity are lacking, you might have to wait a bit before getting a reply.

Citations should refer to this report as: E.B. van Veen, GDPR facts and comments; LIBE and research, MedLawconsult, The Hague, February 2013, and the link on our website.