My health data—your research: some preliminary thoughts on different values in the General Data Protection Regulation

Nikolaus Forgó*

Introduction

Data protection law is about fundamental rights and fundamental values. These values do not necessarily coincide and the fundamental rights reflecting these values quite frequently need to be brought into balance.

Data protection as a fundamental right has made a remarkable career in the last 30 years: starting from a position as a rather obscure side problem that was of no interest to anybody outside a rather small community, it has become one of the most intensely debated fundamental rights in the European Union. Its importance is constantly stressed in mainstream media, in political debates, and also in legal education.

It is, however, important to notice that data protection law and privacy issues were one of the very first problems that lawyers dealing with the upcoming phenomenon of computers and law were interested in. A new academic discipline, called legal informatics, was named then, which mainly dealt with legal knowledge representation and privacy. Privacy laws were among the first legal documents those scholars could then work with. Therefore, a lot of the founding principles of (European) privacy law go back to the 1970s and 1980s, when only a few large computers were used by state authorities, the internet was unknown yet and the PC had not started yet its way into the average person's home. Since then, a lot has changed, not so much in the law, but in the technical circumstances the law needs to deal with. This is specifically true in medical research in which a computer-driven revolution has taken place just like in many other natural sciences. The final goal of many of the medical innovations achieved is to personalize treatment of individuals by better understanding the individual reasons for the disease. This requires a lot of computing and a lot of division of work and therefore a lot of data transfer for research as well as for treatment reasons.

Keypoints

- ICT is a key driver of medical research today.
- ICT-driven medical research often deals with better accessability and connectivity of existing (personal) medical data which brings data protection rules into play. It is of utmost importance to distinguish anonymous from personal data here.
- The upcoming reform of European Data Protection Law will have a significant impact on ICT-supported medical research. Key areas of interest are the distinction between personal and anonymous data and articles 81 and 83 of the Commission's proposal for a General Data Protection Regulation.
- The reform will need to find a precise answer to the question of under which circumstances informed consent is needed or not needed in cases in which existing personal data shall be used for ICT-supported clinical research. No clear

One of the outcomes of this bias between old law and new techniques is that data protection laws are constantly seen as complex, unusable, obstructive, etc.—as they need to be applied to circumstances they were not written for. It is therefore a permanent challenge for data protection lawyers involved in medical research projects to explain their importance and to find arguments why the laws are ethically right. For example, a lawyer needs to insist that each processing of personal data (which was rare in the 1980s and 1990s but is absolutely common today) was (and still is) seen as a potential interference with this person's fundamental rights and therefore per se forbidden and needs justification—by consent or another legal basis.

* Nikolaus Forgó is with Leibniz Universität Hannover, Institute for Legal Informatics, Königsworther Platz 1, D-30167 Hannover, Germany. E-mail: forgo@iri.uni-hannover.de.

Values protected by data protection law run into conflicts with other legally protected values. From open access to the right to be forgotten, from open government to linked data to big data: privacy issues are everywhere and quite frequently data protection lawyers tend to give answers to these issues that are not liked by technical developers. In medicine, the issue is worse, because data protection risks being seen as a hindering factor for the development and exploitation of new knowledge in the patient's best interest.

It is therefore important to remember that data protection as a fundamental right does not come without limits and it does not come without costs. Its very purpose (as with other fundamental rights protecting the patient such as the right to the integrity of the person [Article 3 of the Charter of Fundamental Rights of the European Union]) is challenged by potentially conflicting fundamental rights such as the right to liberty and security (Article 6), the right to freedom of expression and information (Article 11), the right to good administration (Article 41), the right of access to documents (Article 41), etc. As strengthening the fundamental rights to privacy and to data protection can have an impact on these conflicting fundamental rights, these have to be brought into balance. The ECJ puts it very precisely in its recent and already quite famous 'Right to be forgotten' case[1]: a 'fair balance' between conflicting fundamental rights needs to be achieved.[2]

Making the balance 'fair' is an exercise that (constitutional) courts are used to undertaking. The decisions the CJEU has made on data protection-related issues show as well as those of national constitutional courts the need and the ability to weigh conflicting constitutional rights. These judicial attempts are not the focus of this article.

What is of interest here is the fact that prior to the courts already the (European) legislator has to bring conflicting fundamental rights into balance and that bringing these rights into balance needs an understanding of the values that are reflected in these rights.

It is astonishing how unclear these values and their relationships become right from the very moment one tries to look closer at some of the very fundamental provisions of the data protection laws. Writing on these conflicting values would fill a book.

I will therefore choose just one example: the processing of personal data for medical research as it is seen by the (draft) data protection regulation in its different versions. This article therefore does not deal with privacy regulation of medical research in all possible scenarios. When medical research is undertaken within an interventional clinical trial, general data protection laws are complemented by more specific rules, in particular the Clinical Trials Directive (2001/20/EC) and its successor, Regulation No. 536/2014 of the European Parliament and of the Council on clinical trials on medicinal products for human use, which repeals Directive 2001/20/EC and will have to be applied from 28 May 2016. Directive and Regulation both rely on international conventions enshrined in standards of Good Clinical Practice, which are also not a subject of this paper.[3] The complex relationship between these areas of the law remains subject to further consideration. Article 56 par 1 of Regulation 536/2014 will need to serve as a starting point here but will not solve the issue as its meaning remains hard to determine.[4]

However, due to technological developments outside the medical domain, more and more medical data are becoming (theoretically) available and might (theoretically) be used to answer new research questions in the interest of patients outside of (new) interventional clinical trials. ICT for health is an important target in Europe's Digital Agenda[5] and is discussed in a significant number of research projects[6] aiming at improving healthcare via ICT. Many of these projects would benefit from secondary use of data and/or purely observational studies and/or ICT-based evaluations of medical hypotheses.

It is evident that in scenarios in which personal data should be used to answer medical research questions the value of data protection (that shall not be underestimated here) runs into conflict with other values and fundamental rights such as the right to the integrity of the person and the right to freedom of research. One might

1 Judgment of the Court (Grand Chamber), 13 May 2014, C-131/12.

2 Judgment of the Court (Grand Chamber), 13 May 2014, C-131/12, paragraph 81: 'However, inasmuch as the removal of links from the list of results could, depending on the information at issue, have effects upon the legitimate interest of internet users potentially interested in having access to that information, in situations such as that at issue in the main proceedings a fair balance should be sought in particular between that interest and the data subject's fundamental rights under Articles 7 and 8 of the Charter'.

3 It is, however, worth mentioning that general data protection laws and more specific clinical trial regulations do not complement each other smoothly so that the situation becomes significantly more complex when both sets of rules apply. For example, the requirements of informed consent in clinical trials deviate from those in general data protection, see Article 29 par. 1 Regulation 536/2014. Further details on this follow below.

4 '1. All clinical trial information shall be recorded, processed, handled, and stored by the sponsor or investigator, as applicable, in such a way that it can be accurately reported, interpreted and verified while the confidentiality of records and the personal data of the subjects remain protected in accordance with the applicable law on personal data protection'.

5 http://ec.europa.eu/digital-agenda/en/news/ehealth-projects-research-and-innovation-field-ict-health-and-wellbeing-overview.

6 http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=2852.

say that in cases where research leads to tangible results there is a legal and an ethical need to be able to reidentify every patient to let her (potentially) profit from new findings, so that anonymity is not too desirable. One might build up an argument that a patient profiting by better treatment from research undertaken with data that were given to researchers from other patients before he/she developed the disease is under a kind of moral obligation to do the same. Simultaneously, one might counter that our law reserves the right to act unreasonably (within limits), so that such a moral obligation may perfectly well be ignored. It would then possibly be feasible to say that the patient would need to be treated as if she had agreed with the sharing as long as she has not actively rejected it. One might also tend to believe the many affirmations given by physicians that a data subject's attitude towards the sharing of data significantly changes in the moment of a serious diagnosis, in particular if the subject believes that sharing information might have a positive impact on her own outcome. Websites like www.patientslikeme.com or the recently published NIH policy on genomic data sharing[7] show that (some) patients are rather radical in their willingness to share disease-related information. Others are very strict in not sharing anything with anybody.

Big questions arise here, though: How to protect the liberal ideas of autonomy, integrity, and free choice; how to consider public interests; how to allow researchers to make use of their fundamental freedom rights; how to allow progress without jeopardizing patients' interests and so on. I will not be able to give an answer to these questions (neither here nor elsewhere, as I do not have any final answer). What I can do here, however, is to ask how a general data protection rule in abstracto and the draft regulation presented in 2012 and discussed since then in concreto try to balance out these conflicting values and attitudes.

It might be useful to have a closer look at the existing situation today that serves as a baseline.

Directive 95/46/EC

The current debate on data protection reform (obviously) does not start from scratch. Many of the principles are not new but are already part of the existing data protection regime which is mainly based on Directive 95/46/EC. This is, for example, true for the very concept of personal data as such.

However, the data protection regulation draft's ambition was very high. The regulation should, in the view of its authors, 'update and modernize the principles enshrined in the 1995 Data Protection Directive to guarantee the right of personal data protection in the future. They focus on: reinforcing individuals' rights; strengthening the EU internal market; ensuring a high level of data protection in all areas, including police and criminal justice cooperation; ensuring proper enforcement of the rules; and setting global data-protection standards'.[8] A 'comprehensive and coherent approach' should guarantee 'that the fundamental right to data protection for individuals is fully respected within the EU and beyond'.[9]

It is trivial to say that informed consent has been and still is one of the basic principles of European data protection law. The reason is easy to see: informed consent is nothing but a reflection of the old Roman law principle of 'Volenti non fit iniuria': a healthy, well informed individual has the right—and the duty—to decide autonomously how to make use of his rights; she is free to acquire rights, to use them, and to forgo them. Therefore, deciding whether and how one's own data shall be used by others can be seen as just another form of organizing one's personal and professional life via autonomous decisions. Since the famous German census-decision in 1983[10] at the very latest, informed consent and the underlying idea of (informational) self-determination have served as the—probably—most prominent legal basis for processing personal data.

In Directive 95/46/EC informed consent is defined as 'any freely given specific and informed indication of [the data subject's] wishes by which the data subject signifies his agreement to personal data relating to him being processed'.

Article 7 enumerates a list of exceptions in which the processing of personal data is (as an exception) legal. The ground first mentioned in this list is that 'the data subject has unambiguously given his consent' (Article 7a Directive 95/46/EC). When it comes to sensitive data (of which health data are an important example) whose processing is determined in Article 8, again (this time explicit) informed consent is mentioned first (Article 8a Directive 95/46/EC).

Interestingly, Article 8a allows the member states to foresee cases in which explicit consent does not suffice to

7 http://www.nih.gov/news/health/aug2014/od-27.htm.

8 Why do we need an EU data protection reform? http://ec.europa.eu/justice/data-protection/document/review2012/factsheets/1_en.pdf, 2.

9 Communication from the Commission, A comprehensive approach on personal data protection in the European Union, COM(2010) 609 final, 4.

10 http://www.servat.unibe.ch/dfr/bv065001.html; unofficial translation (of the most important parts) via https://www.freiheitsfoo.de/files/2013/10/Census-Act.pdf.

make the processing legal: quite opposite to the Roman 'liberal' idea of self-determination, Member States can define circumstances in which they do not care about the person's will to steer her data processing, but replace or override this will by an 'objective' and collectivistic rule. The paternalistic idea that sometimes the data subject needs to be protected against her own will might be the likely reason for this idea. The values leading to this decision are not clear and it also remains unclear why the decision of whether such a paternalistic approach can be chosen or not is transferred to the level of member state law.

However, another idea which is more important in European day-to-day practice is also enshrined in Articles 7 and 8 of Directive 95/46/EC—which is that there might be cases in which informational self-determination reaches its limits in the sense that informed consent is not needed due to predominant other interests—of the controller and/or of society. Article 7f provides—already on the level of European law—that no consent is needed if processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection [. . .]'.

This is an important entry point for allowing data processing (at least in situations governed by private law) in which the controller cannot or does not want to ask for informed consent: together with Article 7b (which allows processing of data that are needed for the preparation or the performance of a contract) this clause leads to the rather paradoxical situation that processing of personal data by private controllers—opposite to what laypersons believe and opposite to the concept that processing is illegal if there is no exception—often does not need informed consent but simply a good argument why the processing is necessary and not too intrusive instead.

The issue is more complex and more restrictively handled when it comes to sensitive data. No such rule like Article 7f is available here. The principle that (explicit) informed consent is needed prevails. However, an important exception is also made here: member states are allowed to use an exception offered by the European legislator: member states may, for reasons of substantial public interest, lay down exemptions in addition to those laid down in paragraph 2, either by national law or by decision of the supervisory authority, provided that suitable safeguards are in place (Article 8 par. 4). In such cases no informed consent is needed. One of the most commonly used reasons of public interest that leads to such an exception in many member states' legislation is (medical) research. As recital 34 puts it:

> Member States must also be authorized, when justified by grounds of important public interest, to derogate from the prohibition on processing sensitive categories of data where important reasons of public interest so justify in areas such as [. . .] scientific research [. . .].

The outcome of this legislative framework has been problematic in many ways when it comes to international medical research in Europe, in particular when retrospective data shall be used for new research.

First, one has to note that the issue when data are to be seen as anonymous with the consequence that data protection rules do not apply at all is still very unclear and heavily disputed. Data protection law is (probably) the only legal field in which already the very first question of what it should and what it should not regulate is under constant debate. After 30 years of data protection law it is still hard to say when and under which conditions data may be seen as anonymous. Not only are the definitions rather vague (see Article 2a and recital 26), also their interpretation differs significantly within Europe. At the same time, the term 'personal data' serves as an entry point for the applicability of the legal field as a whole. Where there is no personal data there is no legal obligation of data protection. Anonymous data, therefore, need to be distinguished from pseudonymous data,[11] as the former is outside the scope whereas the latter fully falls under the directive's regime. This approach, that uses a very abstract and at the same time very fundamental distinction, is different from legislation in other areas of the world such as the USA where a more pragmatic approach is taken: The Health Insurance Portability and Accountability Act of 1996 (HIPAA), for example, defines, on the one hand, individually identifiable health information and, on the other hand, provides a list of 18 precisely named identifiers that shall be removed in order to achieve de-identified data. There are no restrictions on the use of de-identified data. If the 18 identifiers are removed, this (purportedly) signifies a negligible risk that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is the subject of the information. This list includes values and entries such as names, dates, phone numbers, fax numbers, and e-mail addresses and also URLs, IP

11 A technical concept of pseudonymisation and its implications may be found in ISO/TS 25237:2008 Health informatics—Pseudonymization.

addresses, and full-face photographs.[12] As the list is exhaustive and easily understandable, it is easy to determine whether data can be seen as de-identified or not.

Research projects trying to work in the European environment, which is much more complex to understand and where no list of identifiers that need to be deleted exist, are currently (in my subjective assessment) at risk of following one out of two contradictory but equally false strategies. The first is to underestimate issues of reidentification and reidentifiability and to believe that simply deleting personal identifiers such as name or date of birth is sufficient to render data anonymous. It is very clear that this is not sufficient, as reidentification becomes more and more likely if data are stored and shared and computed for different purposes in different scenarios. Concepts such as de facto anonymity,[13] k-anonymity[14], or datacubes[15] are presented as possible solutions of this issue, but have in common that they are technically and legally sophisticated and require organizational safeguards that need to be adopted from the project's beginning. It is a constant and never-ending challenge to persuade medical researchers that following such rules is a legal requirement that may not be ignored. It is also an ongoing task to enforce the rules and to convince regulators, reviewers, and the scientific audience of their validity. In addition, it needs to be taken into account that in many cases pure anonymization (meaning that reidentification is impossible for anybody) is not a proper way to go, as reidentification might be an ethical and/or legal requirement. This can be the case if it turns out that research undertaken causes unforeseen adverse reactions on the participant's side or produces (incidental) findings with relevance to the participant. In such cases it might be necessary for the physician as an investigator to re-identify the trial participant. This has to be taken into consideration and assessed when setting up a research-environment built on anonymous data.

The alternative error that can be seen in international medical research projects is that the possibility of anonymizing data before sharing it is ignored due to its complexity. This is neither in the interest of the research project nor in the interest of the data subject providing the data. Processing anonymous data should at any rate be the first option to choose if available, but, as always, the devil lies in the detail: Can an MRI image taken in hospital A on day B showing condition C in a patient of age D and having the outcome E be anonymous? Can a gene expression of a patient ever be anonymous, even provided that all matching tables leading to this patient are deleted? Can free-text data in a hospital database ever be anonymized automatically, using methods of text recognition? That nobody can give final answers to these questions is often used as a perfect excuse for not taking the trouble to attempt to rely on anonymous data.

If processing anonymous data is not (or is not believed to be) a suitable option, so that data protection rules apply, it is then just natural to try to legitimatize the processing by informed consent. However, many (retrospective) clinical research projects want to use thousands of patient datasets so that asking all these patients (once more) for informed consent is not an option due to costs and organizational effort. Many of the patients have moved, are deceased, or, more importantly, do not want to be confronted with their disease again so that recontacting them might even be ethically unjustified. Recontacting them would in any case require significant amounts of time and money, and it might also have an impact on the study design and its outcome as it might happen that an insufficient amount of participants can be contacted or that the decision of certain subgroups to participate (or not to participate) again might already have an impact on the composition of the patient cohort.

This issue of reconsent could be organized better if patients were in the position to manage their consent electronically on a platform or via a device that would make it easier for them to know what they have consented to and who wants them to give additional consent. In an ideal world, researchers could easily contact their (former) study participants (if they had consented to be contacted again) on whether they would (again) agree to the processing of their data for another research question. However, not only has such a platform not been developed yet, even if it was technically and economically feasible to offer such a service, it would require changes in the legal system. Electronic consent does not meet per se the legal standards of informed consent in interventional clinical trials.

Thus, clinical trials directive 2001/20/EC defined informed consent (with participation in a clinical trial)

12 See, for example, http://www.hhs.gov/ocr/privacy/hipaa/understanding/summary/privacysummary.pdf; http://www.hipaa.com/2009/09/hipaa-protected-health-information-what-does-phi-include/.

13 Described, for example, in Forgó/Kollek/Arning/Kruegel/Petersen, Ethical and Legal Requirements for Transnational Genetic Research, München 2010. Incorporated in research projects like p-medicine (http://www.p-medicine.eu/), EURECA (http://eurecaproject.eu/), and CHIC (http://chic-project.eu/).

14 L Sweeney, k-anonymity: a model for protecting privacy. (2002) 10 Int J Uncertain Fuzz 557–570, http://arbor.ee.ntu.edu.tw/archive/ppdm/Anonymity/SweeneyKA02.pdf.

15 Described, for example, in E Kamateri and others, 'The Linked Data Access Control Framework' (2014) 50 J Biomed Inform 213–225. doi:10.1016/j.jbi.2014.03.002.

as a 'decision, which must be written, dated, and signed, to take part in a clinical trial, taken freely after being duly informed of its nature, significance, implications, and risks and appropriately documented, by any person capable of giving consent' (Article 2j) Directive 2001/20/EC).

Just like the data protection directive, the clinical trials directive was also under review recently. As well as in data protection, the regulatory approach chosen is no longer a directive but a regulation. Regulation 536/2014 repeals directive 2001/20/EC. The opportunity to allow electronic consent here was missed. The definition of informed consent is still very much the same, is still subject to national law, and still needs to be 'written, dated, and signed' (Article 29 par. 1 Regulation 536/2014).

Projects tend to get stuck at this moment, in particular when they want to make use of retrospective data for new research purposes as it turns out to be impossible to rely on existing informed consent or to go back to (thousands of) patients to ask for reconsent. A way out quite frequently chosen by those not following an Ignore-Data-Protection-Law-Policy is to try to make use of Article 8 par. 4's research exemption. However, as this norm just provides a possibility for member states to create such an exemption in their national system, it is unforeseeable at the time when a (retrospective) study or trial is designed whether all countries where patient data will come from have such an exemption and whether they are compatible with each other. In any case, the legal rules applicable will be very diverse, not harmonized, and very complex to handle.

The outcome of this situation is that medical research projects in Europe, in particular when it comes to research that requires ICT usage and/or exchange of medical data between data controllers located in different European member states and/or relies on retrospective data, are at risk of failing. Researchers tend to underestimate the complexity and importance of these regulatory issues when designing their study, trial, or project and often need to learn at a relatively late stage that the expected data sharing will not happen for reasons of data protection. This is a very costly learning exercise that is not in patients' interest.

Every medical researcher in Europe working in European projects—as well as every patient caring about the issue—will therefore welcome a new data protection regime in the hope that it will ease and clarify the rules for exchange of medical data within Europe. The current rather confusing—not to say: chaotic—situation (in general, not only in medical research) is one of the main reasons why the attempt to choose a regulation instead of a directive as a regulatory instrument was undertaken.[16] Of particular importance for the new framework's usefulness would be (i) a clear definition of personal data allowing an educated guess whether patient data fall under the scope of the (new) rules, (ii) easing informed consent procedures that allow patients to actively manage their consent which require electronic tools, (iii) a clear and European-wide rule when medical research can be undertaken without patient's consent due to overwhelming other values (such as freedom of research or public interest).[17]

The article will now describe the draft regulation's evolution from the Commission's Draft to the Albrecht Draft Report to the Parliament's Frist Reading.

The Commission's draft

In many aspects, the Commission's Draft Regulation (CDR) is an extrapolation of the Directive's existing rules. This is, at the first place, already true for the very concept of personal data itself. Just like the Directive, also the CDR follows a black/white approach, meaning that data are either personal or not. If it is personal, then all data protection rules apply; if it is not, it is outside of the scope of CDR. What personal data are is (again) defined in Article 2, indicating that personal data mean any information relating to a data subject and that a data subject is an identified or identifiable person. The latter is a person:

> [W]ho can be identified, directly or indirectly, by means reasonably likely to be used by the controller or by any other natural or legal person, in particular by reference to an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person (Article 4 par. 1 CDR).

The hint that means reasonably likely to be used need to be taken into account stems from recital 26 of Directive 95/46/EC but was promoted into the main text. Recital 23—similar to the directive's recital 26—amplifies (and repeats) when a person is identifiable:

> To determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used

16 See for example 'How will the EU's data protection reform simplify the existing rules?', http://ec.europa.eu/justice/data-protection/document/review2012/factsheets/6_en.pdf, 'A single set of rules at EU level will have a significant impact on business and enhance the attractiveness of Europe as a location to do business, at the same time as strengthening the EU in its global promotion of high data protection standards'.

17 And, finally, a better coherence between rules on clinical trials and on data protection which is out of scope of this article.

> either by the controller or by any other person to identify the individual. The principles of data protection should not apply to data rendered anonymous in such a way that the data subject is no longer identifiable.

Recital 24 adds some more ambiguity to this by stating 'It follows that identification numbers, location data, online identifiers or other specific factors as such need not necessarily be considered as personal data in all circumstances'. Pseudonymized data would—as today[18]—be treated as personal data with full applicability of data protection rules in so far as there are means likely reasonably to be used either by the controller or by any other person to identify the individual.

So again, it remains unclear whether data that are used for medical research can be anonymous at all as a precondition for the CDR's non-applicability. The underlying difficulty remains that assessing whether the 'data subject is no longer identifiable' is a contextual, fact-based judgement that needs to be taken on a case-by-case basis and which will be very hard to make in many medical scenarios of data sharing. It will be even harder to set up European-wide recommendations or strategies how research projects (in general) should act to achieve anonymity of the data they are processing. In addition, there might be legal and ethical reasons in interventional trials not allowing full anonymization. Chances are very high that—as no legally binding general rule and possibly not even best practices or a code of conduct will be available—today's outcome will be repeated, which is in many cases an abandonment of any attempt to achieve anonymization.

Consequently, again, the concept of informed consent becomes important. The draft data protection regulation repeats the principle that informed consent is the major instrument to allow processing of personal data. Recital 31 of the Commission's draft reads like this:

> In order for processing to be lawful, personal data should be processed on the basis of the consent of the person concerned or some other legitimate basis, laid down by law, either in this Regulation or in other Union or Member State law as referred to in this Regulation.

Informed consent is defined as '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' (Article 4 par 8 CDR). The requirements are tightened here in comparison with the directive as the latter only asks for 'any freely given specific and informed indication' of the data subject's wishes (Article 2h Directive 95/46/EC) so that the term 'explicit' is missing. Further complexity is added here as Article 7 par. 4 CDR adds a new requirement for the validity of informed consent: 'Consent shall not provide a legal basis for the processing, where there is a significant imbalance between the position of the data subject and the controller'.[19] It is not very difficult to say that patients being diagnosed and at the same time asked for consent to the processing of their personal data for research purposes might be in such a position of significant imbalance so that their consent might not be a sufficient basis for the processing.

The CDR does not say anything specific on electronic consent management systems although there is some orientation about written consent given in Article 7 par. 2.[20] As mentioned before, the Clinical Trials Regulation still requires written consent. In addition, it adds some more matters of interpretation as Article 28 of this regulation tackles the issue of broad vs. narrow consent in a rather difficult way:

> *Without prejudice to Directive 95/46/EC*, the sponsor may ask the subject or, where the subject is not able to give informed consent, his or her legally designated representative at the time when the subject or the legally designated representative gives his or her informed consent to participate in the clinical trial to *consent to the use of his or her data outside the protocol of the clinical trial exclusively for scientific purposes*. That consent may be withdrawn at any time by the subject or his or her legally designated representative. *The scientific research making use of the data outside the protocol of the clinical trial shall be conducted in accordance with the applicable law on data protection*.

This allows consent-based research different from the original trial protocol but only if the consent is also valid under a data protection perspective which—again—stresses issues of broad, tied and narrow consent.

These issues of informed consent would be less critcal if the CDR gave a clear answer to the question when research using personal data is to be allowed without consent. Unfortunately, this answer is not clearly given.

18 See in particular Articles 29 WP, WP 136, Opinion 4/2007 on the concept of personal data, 18.

19 Quite similar recital 33: '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.' and recital 34: 'Consent should not provide a valid legal ground for the processing of personal data, where there is a clear imbalance between the data subject and the controller. This is especially the case where the data subject is in a situation of dependence from the controller, among others, where personal data are processed by the employer of employees' personal data in the employment context'.

20 'If the data subject's consent is to be given in the context of a written declaration which also concerns another matter, the requirement to give consent must be presented distinguishable in its appearance from this other matter'. See also recital 32.

Seemingly, the, general, CDR gives orientation on this, specific, situation by providing two articles: one is on 'Processing of personal data concerning health' (Article 81) and the other is on 'Processing for historical, statistical and scientific research purposes' (Article 83). The relationship between those two articles is tricky as Article 81 par. 2 gives a health-specific variant of the general rules on processing for research purposes; however, the content of this specific rule is, roughly, nothing more than a pointer to those general rules of Article 83.[21] Article 83 then draws a, new, line between research as such, which is regulated in Article 83 par. 1, and publication of research which is subject to Article 83 par. 2.

According to par. 1, data may be processed (without consent) if (i) the research purpose cannot be achieved by using anonymous data, (ii) the data are properly pseudonymized, and (iii) the matching table is kept separately from the data.

This rule gives, in principle, precedence to general research interests over privacy interests and values of self-determination and autonomy stand behind. The rule of thumb would be that retrospective research with legally processed, existing data would no longer need informed consent. The data subject's (remaining) interests would be protected by pseudonymization only.

Publication of research data is put under a separate regime (Article 83 par. 2). In this case, again, the data subject's consent (or the fact that she has made the data public, Article 83 par. 2c) which needs to meet the general requirements would be the 'normal' way of making the processing legal (Article 83 par. 2a). There remains still a possibility to publish data without the subject's consent, but the requirements are difficult, so that publicly sharing medical data on that basis only would certainly not be generally advisable: the publication of personal data might still be legal, but only if it

> [I]s necessary to present research findings or to facilitate research insofar as the interests or the fundamental rights or freedoms of the data subject do not override these interests' (Article 83 par 2 b).

It becomes evident that Article 83 is an attempt to bring fundamental values and interests—those of privacy and those of research—into balance. However, it also becomes evident that this balance is not the outcome of an 'objective' procedure or of a broad political consent. Further, it is rather hard to think about medical research without publication—in particular in a world in which more and more publications are required to make the raw data accessible for quality control and research policy reasons[22]—so that the distinction between par. 1 and 2 looks rather artificial. It is therefore not surprising that the line between the affected interests was drawn rather differently in the next stage of the legislative process.

The Albrecht draft report (ADR)

The rapporteur for the Parliament preparing the Parliament's decision-making, the German lawyer Jan-Philipp Albrecht, proposed a very different equilibrium between privacy and research interests. This starts already with a change he proposed on the definition of personal data. Leaving the fundamental distinction between personal and non-personal data untouched,[23] he however proposed to let it suffice for data to be personal if it can be used to single out a person. If a person could be singled out, it would then no longer be relevant whether the data subject could be identified by anybody with reasonable means. In addition, he also suggested to make it harder to qualify data as being anonymous, in particular by amendments to recitals 23 and 24.

Albrecht further suggested to clarify that consent should no longer be a justification for processing 'as soon as the processing of personal data is no longer necessary for carrying out the purpose for which they were collected' (Article 7 par 4a ADR[24]), which would have made it much harder to store research data after the ending of a study or project for (potential) reuse.

In addition, he proposed to change Article 81 par. 2 of the CDR in the way that processing health-related data for research purposes should, as a rule, only be permitted 'with the consent of the data subject'. The (simple) justification given is that 'Health data is extremely sensitive and deserves utmost protection'.[25] In exceptional circumstances, processing should be legal without informed consent, but this would again be subject to national law. Albrecht proposed an Article 81 par. 2a with the following wording:

21 'Processing of personal data concerning health which is necessary for historical, statistical or scientific research purposes, such as patient registries set up for improving diagnoses and differentiating between similar types of diseases and preparing studies for therapies, is subject to the conditions and safeguards referred to in Article 83'.

22 See for example European Medicines Agency, Publication and access to clinical-trial data: an inclusive development process, http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/general/general_content_000556.jsp.

23 It's however worth mentioning that Albrecht proposed to introduce a legal definition for a pseudonym in Article 4, par. 2a).

24 http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2f%2fEP%2f%2fNONSGML%2bCOMPARL%2bPE-501.927%2b04%2bDOC%2bPDF%2bV0%2f%2fEN.

25 http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2f%2fEP%2f%2fNONSGML%2bCOMPARL%2bPE-501.927%2b04%2bDOC%2bPDF%2bV0%2f%2fEN, 198.

> *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* [sic!], *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 competent supervisory authority, in accordance with Article 34(1).

This provision would make processing without consent for research purposes again—just like today—a matter of 28 different jurisdictions. By requiring a prior authorization from a supervisory authority, it would have made up for the lack of the data subject's ability to make use of his informational self-determination and his autonomy by an objective, collectivistic control of an independent bystander (who is not allowed to act irrationally).

In addition and in line with the former principles, Albrecht proposed an amendment to Article 83 CDR which deals with research in general. A new par. 1a would state that the processing of sensitive data and data relating to children may, as a rule, only be processed for research purposes with the data subject's consent. Only for exceptional circumstances should member states be allowed to

> provide for exceptions to the requirement of consent for research [. . .] with regard to research that serves an *exceptionally high public interests* [sic!], if that research *cannot possibly be carried out otherwise*.

Parliament's first reading (PFR)

Albrecht is to be admired for his political skill in convincing the European Parliament to agree with an overwhelming majority under his guidance on a joint text in a first reading—after not less than 3133 proposals for amendments that had been formulated by MEPs. The text, though not surprisingly, clearly shows that it is the outcome of a compromise that did not allow too many differentiations.[26]

The PFR accepts some of Albrecht's proposals on defining anonymous data more restrictively, in particular in recital 23 which reads now as follows:

> The principles of data protection should apply to any information concerning an identified or identifiable natural person. To determine whether a person is identifiable, account should be taken of all the means reasonably likely to be used either by the controller or by any other person to identify or single out the individual directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the individual, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, *taking into consideration both available technology at the time of the processing and technological development. The principles of data protection should therefore not apply to anonymous data, which is information that does not relate to an identified or identifiable natural person. This Regulation does therefore not concern the processing of such anonymous data, including for statistical and research purposes*.

It is worth noting that data that were (originally) anonymous can become personal due to technological developments which would, of course, make it necessary to find a legal basis for the processing in the moment this hardly foreseeable switch from non-personal to personal data occurs. This makes it (once more) unattractive to rely on an anonymity-concept when designing a medical research project.

Article 4 with its definitions underwent important changes as well: the clarification of what personal data is now again very much resembles the existing definition under the Directive, saying in Article 4 par. 2:

> 'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, unique identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social or gender identity of that person.

The promotion of the hint that means reasonably likely to be used needs to be taken into account when deciding whether data are personal is revoked here and, again, exiled into the recitals.[27]

Two more definitions were added: one on pseudonymous and one on encrypted data (Article 4 par. 2a and par. 2b). It is, however, rather likely that these will not have an impact on the basic distinction between personal and non-personal data as both categories deal with personal data, and that both new definitions will be of very limited value for any risk-based attempt to distin-

26 See http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0212+0+DOC+XML+V0//EN for a useful synopsis.

27 Recital 23 PFR: '[. . .] To ascertain whether means are reasonably likely to be used to identify the individual, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration both available technology at the time of the processing and technological development. The principles of data protection should therefore not apply to anonymous data, which is information that does not relate to an identified or identifiable natural person. This Regulation does therefore not concern the processing of such anonymous data, including for statistical and research purposes'.

guish in a more sophisticated way between different categories of personal data.

The requirements for informed consent (Article 7) were alleviated; in particular Article 7 par. 4 was significantly changed so that significant imbalance as a factor devalidating informed consent is no longer mentioned.

Articles 81 and 83 that deal with processing of health-related data and with research underwent, again, serious changes.

Article 81 par. 2 introduces, as it was proposed by ADR, the following new principle:

> Processing of personal data concerning health which is necessary for [. . .] 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.

This consent may be given for one or more specific and similar researches (Article 81 par. 1b). It remains, however, unclear when a research is 'similar', and it also needs to be noted that 'the data subject may withdraw the consent at any time' (ibid).

Article 81 par. 2a then allows the member states to provide for exceptions to this consent requirement, but only 'with regard to research that serves a high public interest, if that research cannot possibly be carried out otherwise'.

Unfortunately, it remains undefined which requirements research needs to meet to serve a high public interest and it will also be very hard to anticipate in a concrete case whether the research could not have been carried out otherwise. Finally, and more importantly, this provision would reintroduce the existing regulatory jungle of 28 different member state laws as it is very apparent that member states (and their data protection authorities) will not agree on the precise meaning of the words either. The Commission and the European Data Protection Board will be entitled to add additional noise here by delegated acts (Article 81 par. 3). The individual patient, potentially affected by that research, will—for obvious reasons—not be heard in this process and will only be able to stress her assumptions a posteriori: She would not have heard in advance of any research undertaken with her data without her consent and she would arguably not have anticipated that such research could be undertaken at all due to the (new) guiding principle that consent of the data subject is needed. If she were not in agreement with the processing of her data for this research, she would then have to challenge the assumption that it was compliant with 81 par. 2 (or that Article 81 par. 2 is compliant with primary law, in particular the Charter of Fundamental Rights of the European Union) after its beginning.

Personal liability risks of researchers involved will further be increased by the new wording of Article 83 which now requires that identifying information (such as the name, address etc.) needs to be separated from other information 'under the highest technical standards' and that 'all necessary measures are taken to prevent unwarranted re-identification of the data subjects'. It is noteworthy that this obligation to use 'highest technical standards' clearly goes beyond the general principles of data security as they are formulated in Article 30 PFR. There, the data controller is obliged to ensure a level of security only with regard to 'the state of the art and the costs of [. . .] implementation'.

Conclusion

Sadly, we are still not in the position to tell how the final text of the regulation will look like due to the never-ending delays in the legislative process.[28] At the same time, this period of political debate allows further reflection on the potential outcome of the regulation.

The regulation as it stands after the PFR will hardly do any good to retrospective medical research in Europe and it will not help patients either, for (at least) the following reasons: It will—as under the present directive—be unclear when the regulation is applicable at all, it will be unclear how exactly informed consent will need to look like in order to be valid and it will again be subject to 28 different legislations under which preconditions consent is dispensable. Patients will not be fully supported in their interest of self-determination and will, again, not be able to actively manage the usage of their data for research purposes. Finally, and most importantly, the fundamental conflict of values—on the one hand in particular patient autonomy and right to integrity, on the other public research interest—is not fully solved in an understandable and coherent way on a European level. Solving this conflict will then again be a task for (national) bodies following national rules. As prior experience with the directive shows, there is some plausibility that the regulation's non-resolution will accompany medical research for the next decades and will evoke another area of uncertainty and doubt. There is still time for debate that could further clarify the situation. Let us use it. The matter is important.

doi:10.1093/idpl/ipu028
Advance Access Publication 19 November 2014

28 And also due to permanent changes in the text. It is noteworthy that according to a draft that was leaked in June 2014 and that is supposed to be the then existing baseline for the Council's work, in particular Article 83 would undergo heavy changes again and would be split up in Articles 83a–84c. See http://statewatch.org/news/2014/jul/eu-council-dp-reg-11028-14.pdf for this version.