Medical Research on Pre-Existing Personal Health Data

The GDPR Exemptions in the European Union

Valérie Junod*
Professor at the Faculty of Business and Economics (HEC) of the University of Lausanne
Professor at the Law School of the University of Geneva
Junod, Muhlstein, Lévy & Puder, Geneva

Daria Gorbacheva
MLaw, Law School of the University of Lausanne

Table of Contents

I. Introduction

More and more data are being produced to be analyzed by more and more powerful tools.[1] This "Big Data" trend is viewed as largely inevitable.[2] It is happening at the hands of commercial, non-profit as well as State parties.[3] Hence, 71% of Europeans consider that providing personal data is just an "increasing part of modern life" for which they have no alternative if they want to consume goods and services.[4] Yet 69% of the respondents in this survey are worried that their personal data might be used by companies and authorities for a different purpose rather than the one based upon which the data were initially collected.[5]

What is true in general is also true for medical data. Treatment of patients and reimbursement of medical services generate a huge amount of health data held by health providers and health insurers. These data can be and are further used in medical research (a practice also referred to as "secondary use" or "retrospective research"). The goal is to generate health benefits for patients through better treatments. The number of research projects using already available data is high and likely to rise even higher.[6] Moreover,

* This paper stems from a talk given at the National Taiwan University during my research leave of 2016 in Taipei. The support of the Taiwan Fellowship is gratefully acknowledged.

1 See recitals 5 and 6 of the General Data Protection Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 (hereafter: GDPR). In addition, many people voluntary put an enormous amount of their personal data, including health information, on social networks where privacy safeguards are – or at least used to be – scarce or inexistent.

2 The "blockchain" trend is also cited as one fostering the use and perhaps the sharing of data, including medical data. See, e.g., Amy Maxmen, AI researchers embrace Bitcoin technology to share medical data, Nature March 9, 2018. Blockchain technology can also be put to use to facilitate medical research. See, e.g., Mehdi Benchoufi/Philippe Ravaud, Blockchain technology for improving clinical research quality, 18(1) Trials (2017).

3 Government agencies are becoming increasingly keen on sharing health data, thus exploiting the potential of Big Data for (applied) research purposes. See for example in Switzerland, Swissmedic, "Big Data" et pharmacovigilance: L'essentiel en bref, 18 Swissmedic Vigilance News (May 2017). In the UK, with respect to the failed care.data initiative, see Fiona Godlee, What can we salvage from care.data? 354 BMJ i3907 (2016). Also Article 29 Data Protection Working Party (hereafter: WP29), Statement on the impact of the development of big data on the protection of individuals with regard to the processing of their personal data in the EU, WP 221, adopted on 16 September 2014. Also on Big Data and health research, see John M. Rumbold/Barbara Pierscionek, A critique of the regulation of data science in healthcare research in the European Union, 18:27 BMC Medical Ethics (2017).

4 See 2015 Special Eurobarometer on Data Protection Report 431, p. 6, available at ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_431_en.pdf.

5 Ibid., p. 71.

6 There are no official statistics of authorized research projects. However, in Switzerland, ethics commissions now publish yearly reports. Taking the one of the canton of Vaud, one reads that the number of retrospective projects went from 104 in 2014 to 128 in 2015; this is to be compared with 34 drug clinical trials in 2015. See 2015 annual report, available at http://www.cer-vd.ch/fileadmin/user_upload/documents/Rapport_Activite_20160429_Final_2.pdf. Using the list of announced Swiss medical studies (2015–2017) on the website of Swissethics (http://swissethics.ch/doc/swissethics/active_research_projects_with_EC_approval.pdf), it appears that about a third of all approved studies are retrospective. One of the reasons favoring retrospective studies is that they are much cheaper and faster to conduct than randomized clinical trials. Another reason is that, compared to clinical trials, they may produce results

these projects are increasingly conducted in a collaborative and cross-border manner[7] by research teams located in different countries.[8] Although medical facilities are the most trusted institutions by Europeans[9], patients remain concerned as to potential misuse of their data.[10,11] Indeed, the regulatory requirements for secondary uses are not always clear, whether for patients nor for researchers.[12]

The former[13] text governing data protection in the European Union (EU), i.e. Directive 95/46/EC, included only few provisions related to scientific research.[14] The entry into force of the GDPR – Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data ("General Data Protection Regulation"[15]) – on May 25, 2018[16] has changed and somewhat simplified this regulatory landscape.[17] For most entities handling data, the GDPR has significantly strengthened the applicable rules.

However, compared to other controllers[18] or processors[19] of personal data, researchers fare well under this new Regulation. When they process (broadly

that shed greater light on real-life medical practice. A third reason is that analysis of available data is increasingly viewed as part of necessary quality assurance.

7 One of the reasons is that it allows to pool more data, thus producing more reliable results, especially when the patient population studied is rather small (e.g., patients suffering from rare diseases). Meta-analyses are a form of results pooling particularly praised in medical research. See, e.g., J.P. Ioannidis/J. Lau, Pooling research results: benefits and limitations of meta-analysis, The Joint Commission journal on quality improvement (1999) Sep;25(9), p. 462-9.

8 For example, in the United Kingdom, "more than half of the UK's research output was the result of an international collaboration". The Royal Society, UK research and the European Union, The role of EU regulation and policy in governing UK research (2016).

9 Eurobarometer 431 (Fn. 4), p. 63.

10 See Patil et al., Public Perception of Security and Privacy: Results of the comprehensive analysis of PACT's pan-European Survey. PACT Project Consortium, 2015. https://www.rand.org/pubs/research_reports/RR704.html.), pp. 35, 46–48. In most countries respondents do not want that their health data be accessed by private third parties, including pharmaceutical companies.

11 In a 2016 UK study, the key concern respondents expressed regarding sharing health data for research purposes was the potential for misuse and abuse. See Aitken Mhairi et al., Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMS Medical Ethics BMS Series, 2016. https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-016-0153-x#Fn47_source).

12 See, e.g., United Kingdom's National Data Guardian (NDG), Review of data security, consent and opt-outs, July 2017, p. 23, at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF.

13 The EU Directive (Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data) applied from December 13, 1995 until May 24, 2018.

14 Subject to Article 7(f) of the former Directive, scientific research fell within the scope of data processing for the purposes of the legitimate interests pursued by the controller, where such interests are balanced with fundamental rights and freedoms of the data subject. (Opinion 6/2014 on the notion of legitimate interests of the data controller under Article 7 of the directive). Article 7(e) provided for data processing for the purposes of public interest as well. Article 11(2) set forth an exemption to the obligation to inform the data subject in cases when the data have not been obtained from the data subject. The exemption applied "in particular for processing […] for the purposes of […] scientific research" if complying with the obligation to inform would be impossible, involve[d] a disproportionate effort as well as if information [was] recorded or disclosed subject to an express provision by law. Article 13(2) enabled Member States to restrict right of access (Article 12) when data were processed solely for the purposes of scientific research, subject to legal safeguards. For an overall comparison of the former Directive and the new GDPR, see Christina Tikkinen-Piri et al., EU General Data Protection Regulation: Changes and implications for personal data collecting companies, Computer Law & Security Review 34 p. 134–153 (2018).

15 On the enactment history of the GDPR, see Mahsa Shabani/Pascal Borry, Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation, European Journal of Human Genetics (2017). The GDPR has been described as the most lobbied piece of EU legislation, with up to 3,999 amendments proposed. See Peter Bolger, Background and Introduction to the General Data Protection Regulation, LK Shields (2017), at https://www.lexology.com/library/detail.aspx?g=d7f59709-4362-4155-ab6f-de55af4147a4; Catherine Stupp, Parliament approves privacy rules after record number of amendments, Euractiv (April 2016), at https://www.euractiv.com/section/digital/news/parliament-approves-privacy-rules-after-record-number-of-amendments/. At the same time the GDPR was adopted, the EU enacted Directive 2016/680 on law enforcement (the Police and Justice Directive); even though it deals with data protection, this Directive is not directly relevant here.

16 Processing which began under the Directive had to be put in compliance with the GDPR before May 25, 2018. The GDPR does not contain transitional provisions, except on very limited set of issues (e.g., Articles 46.5, 91 and 96 GDPR).

17 One of the main objectives in replacing the Directive was to ensure greater harmonization by enacting a Regulation automatically applicable in the 28 EU Member States. Indeed, regulations do not need to be transposed. See, e.g., Paul de Hert/Vagelis Papakonstantinou, The new General Data Protection Regulation: Still a sound system for the protection of individuals, Computer Law & Security Review 32 p. 179–194 (216); by the same authors, Computer Law & Security Review 28 (2012). This does not mean however that national implementing provisions are now superfluous. On the contrary, several provisions will require to be further spelled in national laws. Katrin Schaar has mentioned that 70 possible flexibilities mentioned in the GDPR can be spelled out in national laws. See What is important for Data Protection in science in the future, Working Paper Series of the German Council for Social and Economic Data 258, (2016). William Long and Francesca Blythe mention "30 instances where Member States have been given the ability to legislate at a national level". Member States' derogations undermine the GDPR, Privacy laws and Business, United Kingdom report (May 2016).

18 A controller is anyone who "alone or jointly with others, *determines the purposes* and means of the processing of personal data". Article4(7) GDPR [our emphasis].

19 A processor is "a natural or legal person, public authority, agency or other body which processes personal data *on behalf of the controller*". Article 4(8) GDPR [our emphasis]. As per Article 28 GDPR, processors must be chosen carefully by controllers.

speaking) personal data[20] for scientific research purposes, they can be exempted from several provisions of the GDPR,[21] either directly by the GDPR itself or because the Member States are entitled by the GDPR to introduce further exceptions.

The present paper explores in its chapter II the various exceptions that researchers enjoy. Chapter III of the paper critically assesses these exemptions. The focus is primarily on rules establishing or limiting patients' rights. The conclusion in chapter IV lays down eight broad recommendations.

II. Exemptions in Favor of Research as per the GDPR

A. Brief Overview of the GDPR

As a matter of principle under the GDPR, data subjects enjoy broad rights whenever their personal data are being processed. Personal data are defined broadly and include any kind of data that identify directly or indirectly[22] an individual (as opposed to a legal person);[23] pseudonymized data (i.e. reversibly coded data) are held to be identifiable data and hence personal data.[24] Unless data is anonymous or truly anonymized[25] (irreversibly coded)[26], health data[27] are

20 Research on *truly anonymized* data is not subject to the GDPR and can thus be usually conducted freely, unless national law imposes restrictions. Recital 26. However, truly anonymous or anonymized data are becoming extremely rare. The ability to single out and re-identify an individual within a dataset is enough to make the entire data non-anonymous/non-anonymized. Various studies have shown that this ability to single out is becoming more and more available. See, e.g., Latanya Sweeney, Simple demographics often identify people uniquely, Carnegie Mellon University, Data Privacy Working Paper 3 (2000); Arving Narayanan/Vitaly Shmatikov, Robust de-anonymization of large datasets (How to break anonymity of the Netflix prize dataset), University of Texas at Austin, Working paper, (2008). Only when the cost of identifying an individual would require unreasonable means would that conclusion be discarded. Indeed, according to Recital 26: "[t]o determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments."

21 As recital 4 points out, "[t] he right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality".

22 A natural person is said to be identifiable, even indirectly, when it is ultimately possible to ascertain his identity "by reference to an identifier such as […] location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person". Article 4(1) GDPR.

23 Only living individuals are directly protected by the GDPR. The professional position of the individual is indifferent. Hence, even an individual acting in a professional capacity is within the scope of protection of the GDPR. Deceased persons do not benefit from its protections, unless the Member State where they used to be located decides to extend the scope of protection (see Recital 27).

24 Pseudonymization is defined at Article 4(5) and Recitals 26 and 28 of the GDPR ("'pseudonymisation' means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person;"). Contrary to what is sometimes argued under Swiss law, under the GDPR, the key to decode the pseudonymized data can be held by the same entity without the data losing its status of pseudonymized data. The GDPR does not specify which techniques of pseudonymization must be used. In practice, it is very difficult to know beforehand whether a pseudonymization (or a full anonymization) technique will be effective, especially in the long-term. See, e.g., Matthias Stürzer/Günther Karjoth, Werden Patientdaten anonymisiert? Digma 2017 p. 176; Erik Buchmann, Anonymitätsmasse für Personendaten, Digma 2011, p. 166; Günther Karjoth, Sind anonymisierte Daten anonym genug? Digma 2008 p. 8.

25 Several authors argue that the notion of anonymisation in the GDPR lacks clarity in the light of constantly developing technology enabling reidentification of anonymized data, thus leaving anonymised data vulnerable to privacy breaches. See Francus Aldhouse, Anonymisation of personal data – A missed opportunity for the European Commission., Computer Law & Security Review (2014); For genetic information see Dara Hallinan/Michael Friedewald/Paul De Hert, Genetic Data and the Data Protection Regulation: Anonymity, multiple subjects, sensitivity and a prohibitionary logic regarding genetic data?, Computer Law & Security Review (2013).

26 Recital 26. The two notions are equivalent. However, the anonymization of personal data is still a processing step subject to the GDPR. See Shabani/Borry (Fn. 15). Contrary to US law, there is no safe harbor for anonymization making it highly unsure whether precise health data originating from medical files can ever be truly anonymized. On this issue, see Mark Barnes et al., Impact of the European Union's approved General Data Protection Regulation on scientific research and secondary uses of personal data, 15 Medical Research Law & Policy Report 129 (2016); also WP29 Opinion 06/2013 on open data and public sector information ('PSI') reuse of June 2013, chapter VI.

27 Health data are defined at Article 4(15) GDPR ("data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status;") and further specified by Recital 35 ("all data pertaining to the health status of a data subject which reveal information relating to the past, current or future physical or mental health status of the data subject. This includes information about the natural person collected in the course of the registration for, or the provision of, health care services as referred to in Directive 2011/24/EU of the European Parliament and of the Council (1) to that natural person; a number, symbol or particular assigned to a natural person to uniquely identify the natural person for health purposes; information derived from the testing or examination of a body part or bodily substance, including from genetic data and biological samples; and any information on, for example, a disease, disability, disease risk, medical history, clinical treatment or the physiological or biomedical state of the data subject independent of its source, for example from a physician or other health professional, a hospital, a medical device or an in vitro diagnostic test."). See also under the previous Directive the case of Lindqvist, C-101/01, §50. Genetic data are defined at Article 4(13) of the GDPR and at its Recital 34. Whether genetic data also include information acquired by taking the patient's

personal data and within the scope of the GDPR.[28] Such data even fall within the more protected category of "sensitive" data.[29] Processing is also defined broadly, as any handling of personal data (including storage) falls within this notion.[30]
The basic protective framework of the GDPR can be described as follows:

> "Personal data [must] be processed lawfully, fairly and in a transparent manner in relation to the data subject; collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed; accurate and where necessary kept up to date; […] kept in a form which permits identification of data subjects for no longer than is necessary"[31].

One can add to this already extensive list of rights: the right to be forgotten[32] and the right to data portability.[33] Children enjoy somewhat greater protection:[34] information provided to them must be presented in a manner they can easily understand;[35] in addition, they benefit from a facilitated right to request erasure of their data.[36]
All these obligations fall onto data controllers[37] and (to a lesser extent) onto data processors[38] established within the EU[39] and the EEA.[40] However, the GDPR has extended the geographical scope of application of EU law.[41] Data controllers located outside the EU[42] and EEA (e.g., in Switzerland) are nonetheless sub-

medical history (e.g., genetic diseases incurred by family members) is controversial. However, in any case, such information qualifies as data concerning health. See Gauthier Chassang, The impact of the EU general data protection regulation on scientific research, 11 ecancer medical science 709 (2016); Shabani/Borry (Fn. 15).

28 Some authors have argued that encrypted data should be deemed outside the scope of the GDPR, with regards to parties who do not have the encryption keys. See, e.g., Gerald Spindler/Philipp Schmechel, Personal data and encryption in the European General Data Protection Regulation, Jipitec p. 163–177 (2016). This line of reasoning appears however highly doubtful, given that pseudonymized data are explicitly qualified as personal data. The WP29 has written, in the context of cloud computing, that "encryption may significantly contribute to the confidentiality of personal data if implemented correctly, although it does not render personal data irreversibly anonymous". Opinion 05/2012 of July 1, 2012.

29 Article 9.1 GDPR does not use the term "sensitive", but this term is commonly used in the literature. Other sensitive data include "racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, […] genetic data, biometric data […] data concerning a natural person's sex life or sexual orientation". Genetic data is defined as "*personal data* relating to the *inherited or acquired genetic characteristics of a natural person* which give unique information about the physiology or the health of that natural person and which result, in particular, *from an analysis of a biological sample* from the natural person in question" (Article 4.13 and also Recital 34 GDPR; our emphasis). Whether genetic data can ever be viewed as anonymous or anonymized data is subject to debate. See, e.g., Schaar (Fn. 17); Michael Morrison et al., The European General Data Protection Regulation: challenges and considerations for iPSC researchers and biobanks, 12(6) Regenerative Medicine p. 697 (2017); Dara Hallinan et al., (Fn. 25). It is unclear whether any amount of sensitive data within a more general dataset suffices to make the stricter provisions applicable.

30 Article 4(2) GDPR.

31 Chassang (Fn. 27).

32 Article 17 GDPR. Under the former Directive, from its Articles 12 and 14, the EU Court of Justice derived a right to be forgotten in its well-known Google Spain case (C-131/12). This judgment was controversial; see, e.g, Gregory Voss, The Right to be Forgotten in the European Union: Enforcement in the Court of Justice and Amendment to the Proposed General Data Protection Regulation, Journal of Internet Law 18, no 1 (2014), pp. 3–7; Cesare Bartolini/Lawrence Siry, The right to be forgotten in the light of the consent of the data subject, Computer Law & Security Review 32 (2016), pp. 218–237.

33 Recital 68, 73. Article 20 GDPR.

34 Recital 38. Children also benefit from added protection when they use information society services. Under Article 8, if the information society services are offered directly to a child under 16, a consent should be given or authorized by a holder of parental responsibility in order to enable data processing (Member States are free to establish a different minimum age for processing for these purposes in the limits between 13 and 16). This may become relevant for some medical on-line applications such as health tracers. Concerns have been expressed that the text of GDPR does not take into account advanced levels of commercial literacy of adolescents (as opposed to young children) and that the requirement of parental control could cause excessive parental intrusion into children's lives, eventually leading to a violation of their right to privacy. See Eva Lievens/Valerie Verdoodt, Looking for needles in a haystack: Key issues affecting children's rights in the General Data Protection Regulation, Computer Law & Security Review: The International Journal of Technology Law and Practice (2017).

35 Article 12.1 GDPR and Recital 58.

36 When an individual has given consent to data collection and processing as a child without fully understanding the risks and implications of such consent, they have the right to later request such data to be erased, even as an adult. Recital 65. This should have been made clearer in the Articles of the GDPR.

37 Article 24.1 GDPR.

38 Article 28.1 GDPR. Compared with the former Directive, data processors bear more extensive obligations. However, the rights conferred to data subjects (see chapter III GDPR) are to be exercised against the controller.

39 Article 3.1 GDPR. The notion of establishment is to be understood broadly. See also WP29's Update of Opinion 8/2010 on applicable law in light of the CJEU judgment in Google Spain, adopted on 16 December 2015 (WP 179 update; (opinion issued under the former Directive).

40 Article 7(a) of the Main Agreement on the EEA (EEA Agreement). EEA countries are Norway, Iceland and Liechtenstein. The GDPR was incorporated into the EEA Agreement by the EEA Joint Committee in Brussels on 6 July 2018.

41 Compare under the former Directive: Lokke Moerel, The long arm of EU data protection law: Does the Data Protection Directive apply to processing of personal data of EU citizens by websites worldwide? 1(1) International Data Privacy Law p. 28–46 (2001); also WP 179 update (Fn. 39).

42 Deciding whether a data controller is based in the EU is not that simple, since it is not only the headquarters of the company or the institution which are taken into consideration, but possibly any subsidiaries or branches. Moreover, the GDPR may apply when they are two joint controllers sharing the research responsibility, with one of them being located within the EU. See, e.g., Alan Yeomans/Isabelle Abousahl, Preparing for the EU GDPR in clinical and biomedical research, Viedoc (2017), https://www.viedoc.com/site/assets/files/1323/preparing_for_the_eu_gdpr_in_clinical_and_biomedical_research.pdf.

ject to the GDPR[43] if they collect data from individuals located in the EU for the purpose of providing them with goods or services or if they monitor the behavior of such individuals[44] ("targeting criteria"[45]); the same is true for data processors.[46] Hence, medical researchers based solely in Switzerland are not directly concerned by the GDPR since they offer neither goods nor services. However, in some cases, they may be held to be monitoring the behavior of EU-based individuals. Given the lack of the implementing guidelines on this concept of "monitoring", many researchers based abroad may prefer to abide by the GDPR to be "on the safe side" and to be sure to enjoy the exemptions under the GDPR. In that respect, several authors point out that the GDPR is likely to become the "default global standard" anyway.[47]

When personal data are exploited for research purposes, several of the protective principles mentioned above are curtailed.[48] While the articles of the GDPR do not define research, recital 159 indicates that research should be understood broadly and that it includes: "technological development and demonstration, fundamental research, applied research and privately funded research";[49] public health studies are of course part of scientific research. Research using medical or social science registries is implicitly part of (medical or social science) research.[50] Recital 159 adds that "the Union's objective under Article 179(1) TFEU of achieving a European Research Area" should be taken into account.[51] More helpfully, recital 159 mentions that scientific research goes hand in hand with "publication or otherwise disclosure of personal data". It is not clear how essential this criteria should be.[52] Recital 54 defines *public health,* but in a broad manner, as it encompasses "all elements related to health, namely health status, including morbidity and disability, the determinants having an effect on that health status, health care needs, resources allocated to health care, the provision of, and universal access to, health care as well as health care expenditure and financing, and the causes of mortality".

The next two subchapters outline the restrictions to the GDPR rights, whether they are established by the GDPR directly (subchapter II. B) or made available as an option by the GDPR (subchapter II. C). Subchapter II. D enumerates the rights that data subjects retain, while subchapter II. E mentions the option given to Member States to reinforce certain rights.

B. Exemptions Laid out by the GDPR Directly

In the context of research, the GDPR introduces several exceptions that limit the rights of individuals: i) an exception to the principle of purpose limitation; ii) an exception to the principle of specific consent; iii) an exception to the right of information when the data have been initially collected from third parties;[53]

43 When a non-EU based controller or processor is subject to the GDPR, that person must in principle designate a representative based within the EU. Article 27.1 et 2. That representative becomes the data privacy contact point for authorities and data subjects – instead of, or in addition to, the non-EU controller/processor. Article 27.4. This is a new requirement that the former Directive did not impose. See also Manuel Bergamelli, Die Auswirkung der neuen DSGVO auf die Schweiz, Jusletter April 20, 2018, § 27–32.

44 Collecting health data about individuals in the EU can be viewed as a form of monitoring of their behavior as per Article 3.2.(b) GDPR, although this provision is primarily targeting on-line tracking. See Recital 24 ("In order to determine whether a processing activity can be considered to monitor the behaviour of data subjects, it should be ascertained whether natural persons are tracked on the internet including potential subsequent use of personal data processing techniques which consist of profiling a natural person, particularly in order to take decisions concerning her or him or for analysing or predicting her or his personal preferences, behaviours and attitudes").

45 See, e. g., Federal Data Protection and Information Commissoner, Le RGPD et ses conséquences sur la Suisse, p. 4; Bergamelli (Fn. 43).

46 Article 3.2 GDPR.

47 See Huw Beverley-Smith et al, Faegre, Baker & Daniels, The EU General Data Protection Regulation: Practical Implications for U. S. Businesses, (2017), available at https://faegrebd.com/files/131160_GDPR_Guide_A4_V7.pdf.

48 In earlier drafts of the GDPR, researchers did not enjoy such wide discretion, as data subjects could object to the use of their data for research. However, these protective clauses proposed by the E. U. Parliament were later dropped. See Fn. 126 below.

49 Another definition of research is found at Article 1.1.c of the Commission Regulation 1217/2010 on the application of Article 101(3) TU to certain categories of research and development agreements ("'research and development' means the acquisition of know-how relating to products, technologies or processes and the carrying out of theoretical analysis, systematic study or experimentation, including experimental production, technical testing of products or processes, the establishment of the necessary facilities and the obtaining of intellectual property rights for the results."). Even though this Regulation serves competition purposes, it can help to understand better Recital 159.

50 Recital 157 ("By coupling information from registries, researchers can obtain new knowledge of great value [...] On the basis of registries, research results can be enhanced, as they draw on a larger population. [...] Research results obtained through registries provide solid, high-quality knowledge which can provide the basis for the formulation and implementation of knowledge-based policy, improve the quality of life for a number of people and improve the efficiency of social services."). For an example of such research using registries, see Lina S. Mørch et al., Contemporary Hormonal Contraception and the risk of breast cancer, 377(23) New England Journal of Medicine p. 2228 (2017).

51 According to Article 179(1) TFEU, "[t] he Union shall have the objective of strengthening its scientific and technological bases by achieving a European research area in which researchers, scientific knowledge and technology circulate freely, and encouraging it to become more competitive, including in its industry, while promoting all the research activities deemed necessary by virtue of other Chapters of the Treaties."

52 See Efamro, Esomar, GDPR Guidance Note for the Research Sector: Appropriate use of different legal bases under the GDPR p. 19 (June 2017), available at https://www.esomar.org/uploads/public/government-affairs/position-papers/EFAMRO-ESOMAR_GDPR-Guidance-Note_Legal-Choice.pdf.

53 On this issue, see, e. g., Emma Cradock et al., Nobody puts data in a corner? Why a new approach to categorising personal data is required for the obligation to inform? Computer Law & Security Review 33 p. 142–158 (2017).

iv) an exception to the principle of storage limitation; v) an exception to the rules applicable to sensitive data; v) an exception to data portability; vi) facilitated international transfer of data.

i) Pursuant to the GDPR (without further implementing national provisions), research can be conducted on data which were initially collected for a different (e.g., non-research) purpose. More precisely, the principle of *purpose limitation* does not apply when data are reused for research purposes.[54] For example, if patients provided data to their health professionals for treatment purposes (with or without explicit consent), their data can be reused for research by the same professionals or by others without violating the principle of purpose limitation. In other words, they are not required to consent to this further use. Similarly, if data were collected by insurance companies for payment or reimbursement purposes, they can be reused for research without patients' consent.

For this exception of Article 5 to apply, appropriate technical and organizational safeguards measures must be implemented (Article 89.1) to safeguard the (other) rights of the subject. This notion of appropriate technical and organizations safeguards is often used by the GDPR, but without being precisely defined.[55] Only two examples of such safeguards are provided: data minimization and pseudonymization of data. The first refers to collecting as little data as necessary so to reduce possible risks for individuals.[56] In the context of medical research, this is tricky as researchers are always tempted to accumulate and analyze as much information as possible. The second – as previously explained above – implies that identifiable data are being reversibly coded so that researchers using the coded data cannot infer to which individuals they belong.[57] Other safeguards may be decided by researchers themselves or may be specified by Member States. This may lead to considerable heterogeneity in national practices, thus complicating cross-country collaborations.[58]

ii) In principle, when data subjects are asked to consent to the processing of their data, their consent should only cover the processing activities which were specifically announced to them. Particularly when the data are sensitive, their consent should be explicit and informed;[59] this means that the individual should know exactly how and by whom his or her data will be used.[60] However, when processing is for research, this right is limited, as individuals can give a *general broad consent* that covers a large range of research activities.[61] If they give such a consent, they will not know which future research projects will exploit their data. They will not receive further information as these ulterior projects take place. However, the person may also decide to customize her consent and only agree to the use of her data for certain types of research.[62] Whether this limited consent is binding or can still be bypassed under Article 5 is unclear. Since this is a significant issue, clarification is urgently needed.

iii) Data re-used for research is not always collected from data subjects themselves. For example, researchers may want to use data produced directly by health providers or by (private or public) insurance companies. Even though the data concern the patient, the latter did not provide them directly. In principle, under the GDPR, data subjects must be informed when data concerning them are collected from third parties.[63] However, as per Article 14(5)(b) GDPR, when processing is conducted for the purposes of research, researchers (acting as data controller) are dispensed from informing the data subjects if this would be impossible or would require disproportionate effort.[64] This calls for taking into consideration "the number of data subjects, the age of the data and any appropriate safeguards adopted".[65] Neither a precise number nor a number range is provided – which is perhaps unfortunate. The EU Article 29 Data Protection Working Party (WP29) has added

54 Article 5.1(b).

55 Member States and controllers should further specify the required or retained safeguards. See Shabani/Borry, (Fn. 15); Vigdis Kvalheim/Marianne Myhren, New legislation – a unique opportunity for harmonizing the legal framework for research in the Nordic countries (2017), available at http://www.nsd.uib.no/personvernombud/dok/position-paper-new-legislation.pdf.

56 See also Danny Koevoets, The influence of Article 89 GDPR on the use of big data analytics for the purpose of scientific research, master thesis at Tilburg University (2017), p. 23 (data minimization requires "that the processing of personal data is adequate, relevant and limited to what is necessary in relation to the purpose for which they are processed"). As pointed out by this author, the requirement to apply data minimization appears at odds with the other derogations offered to researchers (e.g., nearly-indefinite storage).

57 Pseudonymization is defined as a safeguard measure – apparently regardless of its effectivity.

58 See Koevoets (Fn. 56).

59 Article 9(2)(a) GDPR. The rules that determine whether consent given is valid have been reinforced under the GDPR, as compared to the former Directive. See also WP29, Guidelines on Consent under Regulation 2016/679, adapted but yet to be finalized (November 2017), at http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=50053.

60 Recital 42. Given the high requirements to be met for consent to be valid, it may be that a waiver of consent through the research exception may be necessary even though some form of (insufficient) consent was initially obtained.

61 Recital 33. See however the comments of the WP29 in its Guideline on consent (WP 259, adopted but still to be finalized): "Considering the strict conditions stated by Article 9 GDPR regarding the processing of special categories of data, WP29 notes that when special categories of data are processed, applying the flexible approach of Recital 33 will be subject to a stricter interpretation and requires a high degree of scrutiny. When regarded as a whole, the GDPR cannot be interpreted to allow for a controller to navigate around the key principle of specifying purposes for which consent of the data subject is asked" (p. 28).

62 Recital 33, Article 21(6) GDPR.

63 Article 14(1) GDPR.

64 See also WP29 Guideline on transparency under Regulation 2016/679 (adopted, but yet to be finalized) WP 260, p. 25 to 28.

65 Recital 62.

that "the impossibility or disproportionate effort must be directly connected to the fact" that the data were obtained from third parties, and not directly from the data subject. This exception is also available to data controllers if informing the data subjects would "render impossible or seriously impair the achievement of the objectives of that processing". Does this clause refer to the risk of yielding less reliable research data? The answer is unclear. In such a situation, "appropriate measures to protect the data subject's rights and freedoms and legitimate interests" must be in place. It is suggested that a possible safeguard is "making the information publicly available". Under this exemption, since subjects are not informed, they cannot object, ask for correction or add relevant information.

iv) Article 5(1)(e) GDPR allows research to be conducted on personal data without complying with the *principle of storage limitation*.[66] In other words, researchers can keep their database of patient data as long as they contemplate further possible research uses. There is nothing in the GDPR that specifies how certain or how precise the future research plans must be. If the research necessitates personal data to be analyzed, then the data need not be anonymized. As previously, appropriate technical and organizational measures must be in place – with no further specification of the concept.

v) Under the GDPR, sensitive personal data enjoys added protection as their processing[67] is not allowed except in a limited set of circumstances enumerated in Article 5(2). However, sensitive personal data can be lawfully processed if this is necessary for public interest, including public health, or scientific research; in that case, the limitation of Article 9 does not apply. Appropriate safeguards – once more – must be in place. Moreover, the processing required by the research must be "proportionate to the aim pursued, [it must] respect the essence of the right to data protection and [it must] provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject." How these requirements are to be understood is not explained by the GDPR. In particular, it is not clear to which extent these requirements go beyond the standard proportionality principle.

vi) Article 17 paragraph 3 introduces a further exception regarding the right to be forgotten. In principle, data subjects can request data controllers to permanently delete their personal data when certain conditions are met, notably that processing of their data is no longer necessary as decided based on a balancing of interests. However, if data controller processes or plans to further process the personal data for research, this right is withdrawn. Two requirements must be met: first, appropriate safeguards as per Article 89(1) must be implemented; second, it must be shown that deletion of the data would either render impossible or seriously impair the research.

Certain types of public health research benefit from an even wider exception.[68] This concerns research "for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy". Given the broad formulation of public health, this second exception could also apply to retrospective research. If this second exception applies, it is no longer required to show that complying with the right to be forgotten would "render impossible or seriously impair the achievement of the objectives of that processing".

vii) Subject to Article 20, the right to data portability only arises when data processing is based on consent or a contract. Therefore, where data is processed under the lawful basis of public interest (Article 6(1)(e)) or of legitimate interests pursued by the controller (Article 6(1)(f)), data controllers do not have an obligation to facilitate data portability.[69] In our context, this means that researchers who are invoking the research exception to process data usually do not have to ensure data portability, even though the initial data were collected from patients with their consent. It is, however, suggested to "develop processes to automatically answer portability requests" as a good practice.[70] An example of such a good practice would be data controllers creating and implementing workable mecha-

66 Principle of storage limitation requires that the data are "kept in a form which permits identification of data subjects for no longer than is necessary for the purpose for which the personal data are processed" (Article 5(1)(e)).

67 For biometric data, there is an opinion that already collection of such data, not only its processing, requires enhanced protection as well. See E. J. Kindt, Having yes, using no? About the new legal regime for biometric data, Computer Law & Security Review: the International Journal of Technology Law and Practice (2017).

68 Under Article 9.2.(i), processing of so-called sensitive data is nonetheless allowed if "necessary for *reasons of public interest in the area of public health,* such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices […]"; under Article 17.3(c), processing necessary "for reasons of public interest in the area of public health in accordance with points (h) and (i) of Article 9(2) as well as Article 9(3);" is exempt from Article 17 on the right to be forgotten; under Article 23.1(2), EU or Member State law can restrict subjects' rights under the GDPR for "other important objectives of general public interest of the Union or of a Member State, in particular […] public health". Public health is also mentioned at Recitals 45, 52–54, 65, 73, 112 and 159.

69 See also Rec. 156: "Member States should be authorised to provide, under specific conditions and subject to appropriate safeguards for data subjects, specifications and derogations with regard […] to data portability […] when processing personal data for […] scientific […] research purposes […]".

70 WP29, Guidelines on the right to data portability, p. 8 (last revised 2017).

nisms (such as digital services) enabling data subjects to access their own data in a user-friendly and machine-readable format, modify, delete and transfer it. This may be useful as more electronic patient files are being developed. In the future, patients may have a clear interest in transferring their health data from one repository to another.

viii) Finally, Article 49(1) offers a somewhat facilitated framework for the transfer of data to third countries.[71] Such transfer may occur when data are being stored outside the EU, for example in a US-based cloud.[72] Although it does not specifically mention research, it applies when "the transfer is necessary for important reasons of public interest" (letter d), which can include matters of public health (see recital 112).[73] In that case, personal data can be transferred even to a country whose standards of data protection are not equivalent to those of the EU;[74] subjects cannot object to such a transfer unless their fundamental rights and freedoms are found to be overriding interests. Moreover, Article 49(1) provides for another exception, allowing non-repetitive transfer of limited amount of personal data to third countries for "purposes of compelling legitimate interests pursued by the controller". Scientific research purposes fall within notion,[75] given "the legitimate expectations of society for an increase of knowledge".[76] Nevertheless, in that second situation, the controller is to inform the data subjects, as well as the supervisory authorities, of the transfer.[77]

C. Additional Exemptions to Be Decided by Member States

Based on the GDPR, each Member State is entitled to introduce certain exceptions. As per Article 89(2), it can decide to lay down derogations from Articles 15, 16, 18 and 21, if it reaches the conclusion that such exemptions are necessary to achieve the research purpose.[78] Whether or not such an exemption is granted,[79] appropriate safeguards through technical and organizational measures must be in place. As already mentioned, these safeguards include the principle of data minimization, pseudonymisation or even anonymization, each time to the extent this is feasible in view of the research objectives. Other safeguards such as rigorous rules on access management are also likely to be appropriate.[80] We review these exemptions in turn:

i) An exemption from article 15 GDPR means that research can be conducted without allowing data subjects to ask whether their data are being processed and how (i.e. to receive information about the purpose of the processing, about the data being used, about the recipients, about the duration of storage, about the other rights available). The subjects are also deprived from their right to access their file and to receive a copy.

ii) An exemption from Article 16 means that the subjects lose their right to *ask for rectification* of their inaccurate data and their right to request completion of their incomplete data. In other words, should a patient be informed of research going on using her data (which is not always a requirement), she would not be able to obtain corrections of incorrect data.

iii) An exemption from Article 18 means that data subjects whose data are being used for research no longer enjoy the right to *restrict* the ongoing research processing, effectively stopping the research as to their own data. This right ordinarily applies when a data subject contests the accuracy of her data, when she claims the processing is unlawful or when she has objected to processing. This right serves to secure a (usually) temporary restriction of processing, pending further checks or further decisions.

iv) An exemption from Article 21 means that subjects lose their right to object to the processing of their data. This exception is only available if the research at issue is necessary to achieve a task in the public interest. This right to object normally exists in situation where the lawfulness of processing derives from prevailing public[81] or private interest. However,

71 According to Mark Barnes et al (Fn. 26), the alternatives – obtaining the consent of a data subject to the transfer or entering a contract containing the model clauses – should usually be more attractive to the data controller.

72 See, e.g., Morrison et al., (Fn. 29, at p. 699). Regarding international transfer through clouds, but under the former Directive, see WP29, Opinion 05/2012 of July 1, 2012, chapter 3.5.

73 Recital 112.

74 The notion of "adequate level of protection", that is a level of protection "essentially equivalent" to that guaranteed in the EU, was developed under the former Directive in the case C-362/14 "Schrems". In that case, the level of protection provided under general US laws was not found essentially equivalent of that applied in the EU. On the other hand, the level of protection in Switzerland is deemed adequate and will remain so under the new GDPR, as long as the Commission does not decide otherwise. See Bergamelli (Fn. 43), §24.

75 Data processing for scientific research purposes is an instance of data processing for the purposes of legitimate interests as per Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of [the former] Directive 95/46/EC, p. 28.

76 Recital 113.

77 Ibid.

78 To be more precise, two requirements must be met, but they appear very similar. First, it must show that maintaining the rights of the subjects would "likely […] render impossible or seriously impair the achievement of the specific purposes". Second, the derogation granted must be "necessary for the fulfilment of those purposes".

79 Article 89(1) requires "appropriate safeguards, in accordance with this Regulation", even when Article 89(2) paving the way "for derogations from the rights referred to in Articles 15, 16, 18 and 21" is not applied.

80 See Shabani/Borry (Fn. 15).

81 Processing is lawful when, for example, it is "necessary for the performance of a task carried out in the public interest". It is also lawful when "necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child." Article 6, letter e) and f).

in case of public interest research, Member States can decide to deprive patients from their opt-out right, meaning that they are forced into accepting public interest research.

The GDPR does not say clearly whether these broad exceptions, if introduced, apply to researchers located in that Member State or to data from subjects located in that Member State. However, the second interpretation would lead to impracticable results because – as already mentioned – datasets typically combine data from subjects located in several countries. It is much more practicable to apply the national rules to researchers based in that country.

However, this second interpretation limits even more the protection benefitting data subjects. The free flow of data from one Member State to others would result in the data of patients located in a restrictive State to be nonetheless subject to the more liberal framework of another Member State. Indeed, nothing prevents a hospital from a Member State which has not made use of the exceptions above to transfer a dataset to researchers located in a different State which has availed itself of all the exceptions.

D. Rights Maintained

Although the following rights, which remain applicable, are not rights specifically granted to data subjects, they deserve mention because they confer added protection to these individuals.

i) Data controllers processing sensitive data are usually obliged to designate a data protection officer (DPO).[82] This is the case when these sensitive data are processed on a large scale within the core activities of data controllers or data processors (in other words, the main activities of a data controller or processor cannot be carried out without processing sensitive data).[83] Instituting a DPO is also required when the processing at issue is done by a public authority, for example a public research institution.[84]

Having a DPO constitutes an added safeguard for data subjects. For example, patients whose data are being used for research can address their questions to this prespecified person.[85] The duties of the DPO include: advising the controller or the processor on the GDPR and other relevant EU or national legislation on personal data protection; monitoring data protection compliance; advising on data protection impact assessment as per Article 35; cooperation with supervising authority; and acting as a contact person to supervising authority on the matters of data protection.[86] Yet, the need to hire a DPO has been criticized as potentially increasing the cost of research and adding to the bureaucratic burden.[87]

ii) Data controllers processing sensitive data on a large scale are obliged to conduct a data protection impact assessment (DPIA) prior to processing.[88] DPIAs aim to prevent violation of the rights and freedoms of the data subjects (in our context patients and research subjects). They assess "the impact of envisaged data processing operations on the protection of persona data", notably "the likelihood and the severity of risks for the rights and freedoms of individuals resulting from a processing operation."[89] The information thus generated allows the controller to decide the measures necessary to address and minimize the risks.[90] The GDPR does not dictate how they should be conducted.[91] Where the DPIA concludes that the rights and freedoms of natural persons may be harmed, the data controller is required to consult a supervising authority before processing the data.[92] These requirements too have been decried as overly onerous,[93] but requests to exempt researchers were not granted.[94]

iii) Data controllers must notify data breaches[95] to the competent national supervisory authorities.[96]

82 Article 37, recital 91. This obligation exists independently of the size of the organization, as the former contemplated limit (250 employees) was dropped in the final version of the GDPR. It applies both to controllers and processors. This is a new requirement that the former Directive did not impose. For more details on data protection officers, see WP29 Guidelines on Data Protection Officers ("DPOs") adopted on 13 December 2016, WP 243.

83 For definitions of "core activity" and "large scale processing", see WP29 Guidelines on DPOs (Fn. 82), pp. 7–8.

84 Article 37(1)(a) GDPR;. also WP243 (Fn. 82), p. 6.

85 Articles 13(1)(b) and Article 14(1)(b) GDPR.

86 Article 39(1) GDPR.

87 See, e.g., Morrison et al., (Fn. 29), p. 699.

88 Article 35(3)(c), recitals 84, 90, 91; also Chassang (Fn. 27).

89 Philip Nolan, Conducting a General Data Protection Regulation compliant data protection impact assessment, Practical Law, Data Protection (2017).

90 WP29, Guidelines on Data Protection Impact Assessment (DPIA), April 2017, p. 4.

91 The W29 guideline offers precise advice, including the recommendation to assess processing operations which began before the entry into force of the GDPR, the recommendation to periodically update the DPIA, the recommendation to publish at least part of the results of the DPIA.

92 Article 36(1) GDPR, recitals 84 and 91 to 94. Data processors are to provide assistance, but do not conduct DPIAs. See further Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679 (2017) at http://ec.europa.eu/newsroom/document.cfm?doc_id=44137.

93 The cost of a single DPIA has been estimated by the EU Commission at between € 14,000 and € 149,000, depending on the type of activity. See United Kingdom Ministry of Justice, General Impact Assessment of the draft GDPR, (2012), at https://consult.justice.gov.uk/digital-communications/data-protection-proposals-cfe/results/eu-data-protection-reg-impact-assessment.pdf. Not complying with the requirement to conduct a DPIA will lead to fines, up to € 10 million or 2% of the turnover. See further Nolan et al. (Fn. 89).

94 In a prior draft, it was contemplated to exempt researchers from conducting DPIAs, but this was ultimately not retained. See Koevoets (Fn. 56), p. 19.

95 Breaches are defined at Article 4(12) GDPR as "the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed".

96 Article 4(13), Article 33. Notification is not required if "the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons". Article 33.1. Processors must notify to their controllers breaches occurring within their sphere of activities. Article 33(2).

When the breach "is likely to result in a high risk" to the data subjects, then the latter must also be notified.[97] This is an important new[98] right of data subjects since health data breaches may have particularly dire consequences. Moreover, past reports indicate that health clinics have been targeted by hackers who then have used the information to blackmail clients.[99] The deadline for complying (72 hours for notifying authorities) is viewed as particularly short, thus forcing controllers to plan in advance and adopt a written "breach procedure".[100]

E. Reinforced Safeguards

With respect to health data, genetic data and biometric data (inter alia), Member States have the option of introducing more severe requirements or limitations.[101] The kind of additional requirements which can be contemplated are not described in the GDPR.

A likely safeguard is the requirement that every retrospective research project be pre-approved by an ethics committee, tasked with verifying whether the interests of data subjects are being properly secured. This is a typical requirement for invasive prospective research projects in the European Union, but, even though not mandatory in all other research settings, it is considered good practice for projects using already collected data.

Another safeguard could be a signed commitment by researchers having access to the pseudonymized database not to try to re-identify its patients.[102] Similarly, such databases could be made available only for consultation in specific locations, forbidding its transfer in digital format.[103] This would reduce the likelihood of re-identification, notably by merging datasets. Such safeguards could be reinforced by corresponding penal sanctions.

III. Critical Assessment and Recommendations

A first general remark is that the GDPR is – at least in large parts – badly drafted. It is difficult to understand exactly to which regime is put a given data from a given subject. The GDPR contains 99 articles, accompanied by 173 recitals whose specific content is not always incorporated in the articles, creating considerable legal uncertainty.

It is not known if researchers and Member States will fully exploit the flexibilities offered by the GDPR. But should that happen, the rights of data subjects are clearly giving way to the interest of society in research. This seems to go against international conventions such as the Biomedicine Convention.[104]

In our view, the balance of interests here weigh too much in favor of researchers.[105] Although medical research is certainly in the public interest, it should not fully override patient autonomy and privacy rights. Indeed, if all flexibilities are exploited, it means that personal patient data can be reused without the data subjects being informed, without consent, without the right to opt-out for an indefinite duration, for an unlimited number of research projects, in an unlimited number of countries.

We believe that a better equilibrium between the rights at issue should be reached. The goal should be to safeguard the trust of the population toward science, researchers and the government. The fact that such research usually does not trigger a direct economic harm for patients (e.g., loss of employment, denial of insurance) does not mean that reasonable

97 Article 4(13), Article 34, Recital 85. Article 34.3 lists three exceptions where notice is not required. On the issues: WP29's Guidelines on Personal data breach notification under Regulation 2016/679, adopted on 3 October 2017, WP250rev.01; also Chassang, (Fn. 27).

98 Previously see Commission Regulation (EU) No 611/2013 of 24 June 2013 on the measures applicable to the notification of personal data breaches under Directive 2002/58/EC of the European Parliament and of the Council on privacy and electronic communications.

99 For example, "[h]ackers [in 2017] publicly posted more than 25,000 files and private images stolen from a Lithuanian plastic surgery clinic, including nude and 'before-and-after' photos, after attempting to financially extort the medical facility and its clients". Hackers post plastic surgery clinic's patient files after blackmail campaign, SC Magazine, June 1, 2017, at https://www.scmagazineuk.com/hackers-post-plastic-surgery-clinics-patient-files-after-blackmail-campaign/article/665357/. In the United Kingdom, "[o]ne in eight consumers in England (13 per cent) have had their personal medical information stolen from technology systems, according to results of a new survey from Accenture". Accenture, Press release of April 25, 2017, at https://www.accenture.com/gb-en/company-news-release-healthcare-data-breached. In the United States, it was calculated that "the final total for individuals impacted by [health] breaches last year [2017] was 14,679,461 – considerably less than the 112,107,579 total the previous year." See Largest Healthcare Data Breaches of 2017, at https://www.hipaajournal.com/largest-healthcare-data-breaches-2017/; also Rumbold/Pierscionek (Fn. 3).

100 See Beverley-Smith et al (Fn. 47).

101 Article 9.4 GDPR. Recital 53 adds that these further conditions "should not hamper the free flow of personal data within the Union when those conditions apply to cross-border processing of such data."

102 See Rumbord/Pierscionek (Fn. 3).

103 Various safeguards to reduce the odds of re-identification were proposed by Chris Culnane et al., Health data in an open world: A report on re-identifying patients in the mbs/pbs dataset and the implications for future releases of Australian government data, University of Melbourne (2017), at https://arxiv.org/pdf/1712.05627.

104 Under Article 2 of the Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine (Convention on Human Rights and Biomedicine), of April 4, 1997, "[t]he interests and welfare of the human being shall prevail over the sole interest of society or science." See Article 16 on consent for research. See also point 8 of the Helsinki Declaration and points 11 to 16 of the WMA Declaration of Taipei on Ethical Considerations regarding Health Databases and Biobanks of 2016.

105 Whether the broad exemptions contained in the GDPR for researchers still comply with the constitutional freedoms of citizens as enshrined notably in Articles 7, 8 and 3.2.a of the EU Charter of human rights remains to be seen.

expectations of privacy can be dismissed. Patients communicate freely with their doctors only because they trust their information will be kept strictly confidential. This reasonable and legitimate expectation is breached if further use is allowed quasi-freely provided that it falls within the broad concept of research.[106] Therefore, we formulate the following recommendations to achieve a better balance between the interests of stakeholders:

i) *Definition of research:* Research eligible for the GDPR exemptions should be defined more precisely. The European Data Protection Board[107] (EDPB; until recently known as the Article 29 Data Protection Working Party or WP29) has not yet issued a specific guideline on the topic.[108] In our view, this notion should not include private commercial research.[109] Whereas it can be argued that patients can expect public-sector researchers to use their data for the collective good, such argument carries less weight when the research is done by private-sector companies aiming for profit. Of course, this is not to say that private-sector research does not produce valuable knowledge, but rather to argue that patients cannot legitimately expect their personal data to be processed by such companies without neither their knowledge nor their consent.

Moreover, and for the same reason mentioned above, only research that is promised for publication should benefit from the exemptions of the GDPR. When this is compatible with the protection of data subjects, the research results produced by one research team should be made available to third party researchers, with reasonable compensation for costs ("sharing of raw data").[110] This shows proper deference for the sacrifices asked of research subjects and/or data subjects who are asked to participate in research projects.

We doubt the need to have a separate definition for "public health research" (as now in recital 54). Given the current broad definition of "public health", practically any kind of medical research can fall within the scope of public health research.

ii) *Scope of application:* The *geographic scope* of the GDPR is not all that clear in the case of scientific research. If a researcher in Switzerland obtains access to a database of pseudonymized hospital files gathered by a French medical team, is she automatically subject to the GDPR? She is not truly monitoring the behavior of EU residents, but she is gaining access to intimate details of their lives. If the GDPR does not apply in full, transfer of personal data must take place in accordance with chapter V, in particular Article 45 on transfers on the basis of an adequacy decision and 46 on transfers subject to appropriate safeguards. In order for Switzerland to retain its status as country that "ensures an adequate level of protection" (article 46), it will have to update its (federal and cantonal) laws on protection of personal data so that they essentially match the GDPR; the process has already begun.[111] Alternatively, safeguards as per Article 46 will need to be implemented, but this option carries greater legal uncertainty.

The scope of *national* data protection provisions should also be clarified. It is not immediately apparent whether the geographic decisive criteria to apply national provisions is the location of the lead researcher, the location of the research lead institution, the location of any researchers, the location of the dataset (e. g., in a cloud[112]) or the location of data sub-

106 In the United Kingdom, the National Data Guardian wrote: "But when patients and service users provide their information to a care professional, they cannot be expected to know all the other uses to which it may be put. There are laws to prevent improper disclosure and procedures to ensure that permission for such 'secondary use' is limited, ethical and secure. However, the laws and procedures are difficult for the experts to understand, let alone the patients and service users. It is hard to argue that patients and service users have consented to uses of their personal confidential information that they cannot anticipate, according to procedures that they cannot understand. This issue is particularly troubling for individuals who have strong views about how their information may be used." Review of data security, consent and opt-outs, July 2017, p. 5. at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF. We note however that, under the NDG's proposal, the opt-out would not be open in case of research on (appropriately) anonymized health data. Id. P. 8.

107 Articles 68 to 76 GDPR.

108 Its 2017 Guideline on consent (Fn. 59) does state: "the WP29 considers the notion [research] may not be stretched beyond its common meaning and understands that 'scientific research' in this context means a research project set up in accordance with relevant sector-related methodological and ethical standards".

109 See the discussion of this notion and the recommendations in Koevoets (Fn. 56).

110 See, e. g., Andrew J Vickers, Whose data set is it anyway? Sharing raw data from randomized trials, Trials. 2006; 7: 15; same author, Sharing raw data from clinical trials: what progress since we first asked "Whose data set is it anyway?", Trials. 2016 May 4;17(1):227; Alawi A. Alsheikh-Ali et al., Public Availability of Published Research Data in High-Impact Journals, PLoS ONE 6, e24357 (2011); Florian Naudet et al., Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in The BMJ and PLOS Medicine, BMJ 2018;360:k400; as well as the many articles published by the New England Journal of Medicine and brought together on this page: http://www.nejm.org/data-sharing.

111 See the project fully revising the Swiss Data Protection Act and the Message of the Federal Council of September 15, 2017, FF 2017 p. 6565; the parliamentary debates (object 17.059) are available from https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20170059, The Swiss trade associations are worried that the pace of the Swiss DPA revision might be too slow in view of feared "eurocompatibility" problems. See economiesuisse, Les entreprises suisses ont besoin d'une protection des données sur mesure, Press release of February 1, 2018, at https://www.economiesuisse.ch/fr/articles/les-entreprises-suisses-ont-besoin-dune-protection-des-donnees-sur-mesure.

112 In this respect, see Mark Webber, The GDPR's impact on the cloud service provider as processor, Volume 16(4), Privacy & Data Protection (2016); also EU Data Protection, Code of conduct for cloud service providers (May 2017).

jects. Given the impact that this can have on individual data protection, this ought to be urgently clarified. Similarly, it should be clarified under which conditions storing health data on a cloud or on a foreign-based server is allowed.[113] This is all the more important since an increasing number of computer software or cellphone applications carry nearly automatic safeguards on clouds whose locations are often unknown.[114]

iii) *Pseudonymization techniques:* The GDPR marks a clear shift towards systematic *pseudonymization* of research data (see recitals 28, 29, 78, 156). However, researchers and ethics commissions are too often unsure how pseudonymization needs to be performed on their particular datasets.[115] This is especially true when several databases are pooled or merged, a situation for which no specific rules are currently laid down. Hence, appropriate pseudonymization techniques in the field of medical research should be laid down in EU guidelines.[116] These guidelines should further specify the (estimated) residual risk of reidentification held to be (still) admissible (e.g., 1%, 0.1%, 0.01%?). In our view, it is only once a minimum standard for pseudonymization has been defined, that compliance with the said standard could fairly lead to a more relaxed research framework. Hence, we suggest that the exemptions offered by the GDPR be made dependent on high-quality pseudonymization.

iv) *Transparency:* We recommend that all projects using personal health data without subjects' consent should be reported on a central EU platform so that individuals can get a general understanding of what projects are likely to have used, or to be using, their data. The database should further report on the pseudonymization techniques used for each project. The public debate about the proper (and possibly improper) reuse of personal data would thus be encouraged.

This central database could also inform the public about the choices made by Member States to use the flexibilities offered by the GDPR. This way, patients from various countries could consult one central website to know exactly which options have been retained by each country. As mentioned above, the countries with the least restrictive regime could end up attracting either the researchers or the datasets that stem from patients located in Member States having retained a more protective regime. Having a centralized information access pathway would somewhat compensate for this disadvantage.

There should be provisions to clarify how the data protection requirements, particularly the data minimization, can be reconciled with the recent publicity requirements of funding organisms.[117] In many countries, public funds are granted only if researchers commit to making their research results publicly available upon request; increasingly, scientists are asked to deposit their raw data in public repositories. This is the case of the Swiss National Science Foundation.[118] In principle, this requirement from funding agencies usually applies to pseudonymized datasets. However, putting pseudonymized datasets in the public domain may well lead to further curtailment of the rights of data subjects. Given the impact that these requirements by public funding agencies will have over time, this deserves to be the topic of specific provisions.

v) *External control:* It should be clarified who bears the burden of proving that the requirements laid down by an exemption are met and how this proof must be provided. Reading the GDPR, this proof appears to be specific to each research project; unless a Member States provides otherwise, researchers can self-confirm that their projects meet the requirements. We believe that, in each Member State, an administrative authority (e.g., an ethics commission) should be entrusted with the mission of verifying compliance.[119] Its decision should be subject to appeal, for example by patient organizations or by individuals whose data are likely to be exploited (e.g., the patients of a hospital).

vi) *Lawfulness:* It is not entirely clear which is the legal basis that makes scientific research on existing

113 The GDPR currently does not contain specific provisions regarding to storage of health data on cloud servers. Depending on the particular functions carried out by cloud services, the latter are held to be data controllers or data processors. *See* Marina Škrinjar Vidović, EU Data Protection Reform: Challenges for Cloud Computing, Croatian Yearbook of European Law and Policy, Vol 12 (2016), p. 176). The WP29 has issued recommendations of cloud computing, but under the former Directive; some of its recommendations are nonetheless likely to be still relevant and applicable. In particular, clients "should select a cloud provider that guarantees compliance with EU data protection legislation". The WP also stresses the need for and importance of transparency: WP29, Opinion 05/2012 on cloud computing, WP 196, 01037/12, adopted on July 1, 2012.

114 See, e.g., Bernadette John, Are you ready for General Data Protection Regulation? Editorial, BMJ (2018).

115 See in Switzerland Stürzer/Karjoth (Fn. 24).

116 Compare in the United States with US Department of Health and Human Services, Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, at https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html. See also for the EU: European Medicines Agency, External guidance on the implementation of the EMA policy on the publication of clinical data for medicinal products for human use (September 2017), at http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2017/09/WC500235371.pdf.

117 See, e.g., Morrison et al., (Fn. 29), p. 698.

118 See Swiss National Science Foundation's policy on Open Research Data, available at: http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/default.aspx#SNSF%20policy%20on%20 Open%20Research%20Data.

119 This is the case in Switzerland pursuant to Article 34 of the Federal Law on human research. Based on the 2015 annual report of the Vaud's ethics commission, it appears that about half retrospective projects obtained a consent waiver. See Fn. 6 above.

datasets lawful. It is a general principle under the GDPR that every data processing must be lawful in the sense that it relies on one of the grounds of Article 6. Whereas Article 9.2(j) mentions research as a lawful ground to justify the processing of sensitive data ("special categories of personal data"), this is not the case at Article 6. Briefly described the six available lawful grounds at Article 6 are: (a) free, informed, specific and affirmative consent; (b) necessary performance of valid contract; (c) necessary compliance with a legal obligation of the controller; (d) necessary protection of the vital interests of an individual; (e) necessary performance of a task carried out in the public interest or in the exercise of official authority vested in the controller; (f) legitimate interests of the controller not overridden by the interests of the data subject.[120] In the case of retrospective research using the flexibilities of the GDPR, only grounds (e) and (f) can enter into consideration. Researchers cannot usually rely on consent since they do not fully satisfy the conditions for valid consent (e.g., the subject's consent must be revocable, it must be unbundled and granular; there must be no imbalance of powers).[121] Neither ground (e) nor ground (f) provides ideal lawful basis in the context of medical research: ground (e) because not all research is in the public interest (see already point i) above); ground (f) because it requires to apply the "not overridden by the interests of the data subject" criteria in a situation where the GDPR itself contains so many derogations against data subjects. It would be helpful if the EU authorities would clarify which ground should be relied upon and – above all – how.[122]

vii) *Further non-scientific use of scientific databases:* It is unfortunate that the GDPR does not incorporate a ban on use of health data acquired by illegal (i.e., non-GDPR compliant) means. It would be reassuring to data subjects to know that, for example, insurance companies, employers and banking-credit institutions are banned from using data that have become available "thanks" to a breach of GDPR obligations. Similarly, the GDPR should have banned re-use by government services, especially law enforcement and national security services, of health data repurposed for research.[123] Given that health data can be of great interest to government services to elucidate crime or to fight terrorism, there is an obvious appeal to obtain data from scientists. These sources of data, especially genetic data, are all the more attractive since researchers are now able to merge or pool together gigantic datasets using the flexibilities offered by the GDPR.[124] It would be hardly surprising if, a few years from now, the most complete databases of populations were to be found in the hands of researchers. Clearly forbidding access police and national security agencies to access these databases is a necessity.[125]

viii) *Opt-out:* In our view, an opt-out right should have been maintained as contemplated in one of the prior GDPR drafts.[126] Our conclusion stems from the (increasing) inability to prevent re-identification of datasets, even when pseudonymization or anonymization techniques have been used. Several studies have shown that the odds of being able to correctly single out and identify individuals in a database remain high.[127] If it cannot be faithfully promised with near 100% security that re-identification is

120 See also the ECJ case of *Rīgas* of May 4, 2017, Case C-13/16.

121 See WP29, 2017 Guideline on consent, (Fn. 59), especially p. 22 and 29 ("It is important to remember that if consent is being used as the lawful basis for processing, there must be a possibility for a data subject to withdraw that consent.").

122 WP29 has already stated that: "The application of one of these six bases must be established prior to the processing and in relation to a specific purpose. As a general rule, a processing activity for one specific purpose cannot be based on multiple lawful bases." (p. 22).

123 The GDPR does not apply to use of personal data by competent law enforcement authorities. Article 2.2.(d). Partially on this issue, see the very interesting article by Catherine Jasserand, Law enforcement access to personal data originally collected by private parties: Missing data subjects' safeguards in directive 2016/680, Computer Law and Security Review 34 p. 154–165 (2018). See also in the United Kingdom, Gareth Iacobucci, NHS must stop sharing confidential patient data with Home Office, says MPs, BMJ 360 (2018); more generally WP29, Opinion 01/2012 on the data protection reform proposals, adopted on March 23, 2012.

124 The GDPR contains no provisions specifically on the pooling of data, notably health data.

125 In its Working Document on surveillance of electronic communications for intelligence and national security purposes of December 2014, the WP29 described how internet surveillance had been taking place outside the legal boundaries, making it clear that public agencies cannot always be trusted to comply with international and national rules guaranteeing privacy.

126 In May 2014, the EU Parliament had proposed the following text for Article 81.2a: "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 a *high public interest,* 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 unwarranted re-identification of the data subjects. However, *the data subject shall have the right to object at any time* in accordance with Article 19." See Position of the European Parliament adopted at first reading on 12 March 2014 with a view to the adoption of Regulation (EU) No …/2014 of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data, P7_TC1-COD(2012)0011 (our emphasis). These provisions were later deleted. They were opposed by pro-research groups. See, e.g., Federation of European Academies of Medicine & Wellcome Trust, Realising the societal benefits of health research through the Data Protection Regulation (2012/0011(COD)), at https://wellcome.ac.uk/sites/default/files/briefing-societal-benefits-of-health-research-through-data-protection-regulation-wellcome-feb13.pdf.

127 See, e.g., Culnane et al. (Fn. 103); also Khaled El Emam et al., A systematic review of re-identification attacks on health data, PLOS (2011), at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0028071. More generally on the issue Feng-Jen Tsai/Valérie Junod, Medical research using governments' health claims databases: with or without patients' consent? Journal of Public Health (2018).

impossible,[128] then opt-out should be a minimum choice offered to all patients.[129] As the UK National Data Guardian wrote, "the principle of offering an opt-out is core to building public trust. People need to see they can exercise control and that data is not being used in a way which will surprise them."[130] Exceptions could be granted on a case-by-case basis when proven before an administrative authority that truly important research cannot be otherwise performed. The opt-out procedure should be simple for data subjects and could even be global (e.g., refusing or consenting to any reuse of any medical data by any public institutions).[131] This would be an improvement over the current GDPR regime where patients have neither an opt-in nor an opt-out right. Patient choice would then be entered into a central database, which could be queried by researchers across the EU.

A fortiori, reuse of the health data stemming from incompetent minors should not be authorized without parental consent. A significant amount of information may be collected from minors without them being aware of the situation. Privacy of individuals who do not choose freely to interact with given health providers or insurance companies deserve even higher degree of protection. At their majority, these data subjects should be given the option to formally opt-in or opt-out.

Finally, it ought to be made clear whether the multiple research exemptions described in this paper can still be invoked when data were initially gathered *subject* to certain constraints mutually agreed between the data subjects and the controllers? For example, when the data subject signed a consent form to participate in a drug clinical trial. We believe that if data were initially collected based on certain explicit promises (e.g., no further reuse, no further transmission), these promises should be held, despite the leeway offered by the GDPR.

IV. Conclusion

Given the significant leeway offered by the GDPR, it will be essential to study how this regulation will be implemented in the field of medical research. Ideally, a formal assessment should be conducted, based *inter alia* on surveys of scientists' practices and patients' expectations. The costs of complying with the GDPR should be further appraised. Whether or not this new legislation will simplify the life of researchers remains to be seen.

Probably the greatest added value of the legislation would come from rules that are easy and uniform to both apply and to explain. In that respect, the complexity of the GDPR rules generates significant doubt. This uncertainty is all the more difficult to bear given the high potential fines for GDPR violations.[132]

A broader solution would be to make retrospective medical research the topic of a separate EU regulation with uniform standards across Member States.[133] In the United States, this approach has been criticized,[134] but mainly because U.S. law still lacks a general framework for data privacy. In Europe, specifying the rules for retrospective research would allow researchers and data subjects to know more precisely what they are entitled to do, respectively what they must tolerate, for the sake of science.

128 For example, in the United States, the Safe Harbor Standard under HIPAA (Health Insurance Portability and Accountability Act) has been estimated to carry a risk of only 0.04% of individuals to be uniquely identifiable. See, e.g., Ann Cavoukian/Daniel Castro, Big Data and innovation, Setting the record straight: De-identification does work, p. 5 (2014), at http://www2.itif.org/2014-big-data-deidentification.pdf. See also Khaled El Emam et al., De-identification methods for open health data: the case of the Heritage Health Prize claims dataset, 14(1) Journal of Medical Internet Research e33 (2012)., where the cutoff was 0.05%.

129 In the United Kingdom, after the scandal of the failed care.data project, the government decided that everyone was to be granted "the choice to opt out of sharing their data beyond their direct care". U.K. Department of Health, Your Data: Better Security, Better Choice, Better Care, Government response to the National Data Guardian for Health and Care's Review of Data, Security, Consent and Opt-Outs and Care Quality Commission's Review, 'Safe Data, Safe Care' (July 2017), p. 6.

130 National Data Guardian for Health and Care 2017 report, Impact and influence for patients and service users, 12 December 2017, p. 10.

131 For example, hospitals could have an entry form signed by incoming patients stating whether they wish to opt out from retrospective data. Patients who do not opt-out would be considered "opt-in". The form should be presented at regular intervals (e.g., every 3 years) to patients.

132 Article 83, para. 4.5 ("Infringements of the following provisions shall, in accordance with paragraph 2, be subject to administrative fines up to 10,000,000 EUR, or in the case of an undertaking, up to 2% of the total worldwide annual turnover of the preceding financial year, whichever is higher [...]"). For more details: WP29's Guidelines on the application and setting of administrative fines for the purposes of the Regulation 2016/679, adopted on 3 October 2017, WP 253. See also Koevoets (Fn. 56). The level of fines has been compared to that imposed in case of anticompetitive or corruptive behaviors. See Beverley-Smith et al (Fn. 47).

133 As research is a shared competence of the European Union and Member States, the European Union could choose to enact more detailed regulations in the field of retrospective research. See Royal Society, (Fn. 8).

134 See, e.g., N. Terry, Existential challenges for healthcare data protection in the United States, 3 Ethics, Medicine and Public Health, p. 19–27 (2017).