HEALTHY DATA PROTECTION

*Lothar Determann**

Modern medicine is evolving at a tremendous speed. On a daily basis, we learn about new treatments, drugs, medical devices, and diagnoses. Both established technology companies and start-ups focus on health-related products and services in competition with traditional healthcare businesses. Telemedicine and electronic health records have the potential to improve the effectiveness of treatments significantly. Progress in the medical field depends above all on data, specifically health information. Physicians, researchers, and developers need health information to help patients by improving diagnoses, customizing treatments and finding new cures.

Yet law and policymakers are currently more focused on the fact that health information can also be used to harm individuals. Even after the outbreak of the COVID-19 pandemic (which occurred after the manuscript for this article was largely finalized), the California Attorney General Becera made a point of announcing that he will not delay enforcement of the California Consumer Privacy Act ("CCPA"), which his office estimated imposes a $55 billion cost (approximately 1.8% of California Gross State Product) for initial compliance, not including costs of ongoing compliance, responses to data subject requests, and litigation.†

Risks resulting from health information processing are very real. Contact tracing and quarantines in response to SARS, MERS, and COVID-19 outbreaks curb civil liberties with similar effects to law enforcement investigations, arrests, and imprisonment. Even outside the unusual circumstances of a global pandemic, employers or insurance companies may disfavor individuals with pre-existing health conditions in connections with job offers and promotions as well as coverage and eligibility decisions. Some diseases carry a negative

---

* Lothar Determann teaches computer, Internet, and data privacy law at Freie Univer sität Berlin, University of California, Berkeley School of Law, and Hastings College of the Law, San Francisco, and he practices technology law as a partner at Baker McKenzie in Palo Alto. Opinions expressed in this article are those of the author, and not of his firm, clients, or others. The author is grateful to Arian Grüner for assistance with research and editing from a legal and medical perspective, as well as valuable input from Prof. Dr. Ulrich M. Gassner, Priv-Doz. Dr. Felix Post, and Dr. Jasper zu Putlitz.

† *See* Joe Duball, *California Attorney General's Office: No Delay on CCPA Enforcement Amid COVID-19*, IAPP (Mar. 24, 2020), https://iapp.org/news/a/making-sense-of-calls-to-delay-ccpa-enforcement-amidst-covid-19; *see also* Berkeley Economic Advising and Research, LLC, *Standardized Regulatory Impact Assessment: California Consumer Privacy Act of 2018 Regulations* at 19 (Aug. 2019), www.dof.ca.gov/Forecasting/Economics/Major_Regulations/Major_Regulations_Table/documents/CCPA_Regulations-SRIA-DOF.pdf.

stigma in social circumstances. To reduce the risks of such harms and protect individual dignity, governments around the world regulate the collection, use, and sharing of health information with ever stricter laws.

European countries have generally prohibited the processing of personal data, subject to limited exceptions, for which companies have to identify and then document or apply. The General Data Protection Regulation ("GDPR") that took effect in 2018 confirms and amplifies a rigid regulatory regime that was first introduced in the German State Hessen in 1970 and demands that organizations minimize the amount of data they collect, use, share, and retain. Healthcare and healthtech organizations have struggled to comply with this regime and have found EU data protection laws fundamentally hostile to data-driven progress in medicine.

The United States, on the other hand, has traditionally relied on sector- and harm-specific laws to protect privacy, including data privacy and security rules under the federal Health Insurance Portability and Accountability Act ("HIPAA") and numerous state laws including the Confidentiality of Medical Information Act ("CMIA") in California, which specifically address the collection and use of health information. So long as organizations observe the specific restrictions and prohibitions in sector-specific privacy laws, they may collect, use, and share health information. As a default rule in the United States, businesses are generally permitted to process personal information, including health information. Yet, recently, extremely broad and complex privacy laws have been proposed or enacted in some states, including the California Consumer Privacy Act of 2018 ("CCPA"), which have a potential to render compliance with data privacy laws impractical for most businesses, including those in the healthcare and healthtech sectors.

Meanwhile, the People's Republic of China is encouraging and incentivizing data-driven research and development by Chinese companies, including in the healthcare sector. Data-related legislation is focused on cybersecurity and securing access to data for Chinese government agencies and much less on individual privacy interests.

In Europe and the United States, the political pendulum has swung too far in the direction of ever more rigid data regulation and privacy laws, at the expense of potential benefits through medical progress. This is literally unhealthy. Governments, businesses, and other organizations need to collect, use and share more personal health information, not less. The potential benefits of health data processing far outweigh privacy risks, which can be better tackled by harm-specific laws. If discrimination by employers and insurance companies is a concern, then lawmakers and law enforcement

agencies need to focus on anti-discrimination rules for employers and insurance companies - not prohibit or restrict the processing of personal data, which does not per se harm anyone.

The notion of only allowing data processing under specific conditions leads to a significant hindrance of medical progress by slowing down treatments, referrals, research, and development. It also prevents the use of medical data as a tool for averting dangers for the public good. Data "anonymization" and requirements for specific consent based on overly detailed privacy notices do not protect patient privacy effectively and unnecessarily complicate the processing of health data for medical purposes.

Property rights to personal data offer no solutions. Even if individuals - not companies creating databases - were granted property rights to their own data originally, this would not ultimately benefit individuals. Given that transfer and exclusion rights are at the core of property regimes, data property rights would threaten information freedom and privacy alike: after an individual sells her data, the buyer and new owner could exercise his data property rights to enjoin her and her friends and family from continued use of her personal data. Physicians, researchers, and developers would not benefit either; they would have to deal with property rights in addition to privacy and medical confidentiality requirements.

Instead of overregulating data processing or creating new property rights in data, lawmakers should require and incentivize organizations to earn and maintain the trust of patients and other data subjects and penalize organizations that use data in specifically prohibited ways to harm individuals. Electronic health records, improved notice and consent mechanisms, and clear legal frameworks will promote medical progress, reduce risks of human error, lower costs, and make data processing and sharing more reliable.

We need fewer laws like the GDPR or the CCPA that discourage organizations from collecting, using, retaining, and sharing personal information. Physicians, researchers, developers, drug companies, medical device manufacturers and governments urgently need better and increased access to personal health information. The future of medicine offers enormous opportunities. It depends on trust and healthy data protection. Some degree of data regulation is necessary, but the dose makes the poison. Laws that require or intend to promote the minimization of data collection, use, and sharing may end up killing more patients than hospital germs.

In this article, I promote a view that is decidedly different from that supported by the vast majority of privacy scholars, politicians, the media, and the broader zeitgeist in Europe and the United States. I am arguing for a healthier balance between data access and data

protection needs in the interest of patients' health and privacy. I strive to identify ways to protect health data privacy without excessively hindering healthcare and medical progress. After an introduction (I), I examine current approaches to data protection regulation, privacy law, and the protection of patient confidentiality (II), risks associated with the processing of health data (III), needs to protect patient confidence (IV), risks for healthcare and medical progress (V), and possible solutions (VI). I conclude with an outlook and call for healthier approaches to data protection (VII).

All of Us is the name of an ambitious project launched by the National Institutes of Health ("NIH") in 2017 to collect health data and genetic information from one million U.S. citizens in a national research database.[1] The title conveys a variety of meanings and associations. "All of Us" can mean "everyone" or "everything from us." This ambiguity reflects how hu-

1. NATIONAL INSTITUTE OF HEALTH, ALL OF US RESEARCH PROGRAM OPERATIONAL PROTOCOL 5 (March 28, 2018), https://allofus.nih.gov/sites/default/files/aou_operational_protocol_v1.7_mar_2018.pdf; *see also* Lothar Determann and Felix Post, *Gesunder Datenschutz*, *in* DIE ZUKUNFT DER MEDIZIN 317 (Erwin Böttinger & Jasper zu Putlitz eds., 2019).

man beings have experienced their role in the development and application of new diagnoses, treatments, drugs, and medical devices over time.

I. MEDICINE GETS PERSONAL

"Everything from us" invokes the use of "big data" technologies in medicine. "Big data" is generally understood to mean analyzing particularly large volumes, variety, and velocity of data,[2] and often includes information that was originally collected for different purposes. "Everything from us" implies the increasing collection of health data in the course of medical progress: resident doctors, hospitals, and health insurance companies keep patient files in electronic form; digital health apps accompany us in real time every step of the way.[3] More and more medical devices in hospitals and doctor's offices, which previously offered independent and stand-alone functionalities, are now integrated into networks and deliver large amounts of data to be further processed and stored via clinic information systems ("CIS") and software.[4] So-called "biobanks," which are databases of biological and genetic material, have been established and form the foundation on which digital processing and the exchange of genetic information takes place.[5] At the same time, telemedicine, in which medical consultation and treatment is carried out remotely and online, has grown exponentially, sometimes making the personal interaction between patient and doctor obsolete altogether.[6] The wealth of existing data, including increasing information about the genetic composition of human beings, provides new opportunities for medical research and treatment.[7] Medication and treatment options tailored to the patient's specific life circumstances, environmental

2. The concept has been described as the three Vs by Doug Laney in 2001. *See* Doug Laney, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, META GROUP INC. (Feb. 6, 2001), http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

3. *See, e.g.*, STANFORD MED., STANFORD MEDICINE 2017 HEALTH TRENDS REPORT: HARNESSING THE POWER OF DATA IN HEALTH 15 (2017).

4. For an overview of CIS, see Donte, *What Is A Clinical Information System (CIS)?*, BIOHEALTHMATICS (June 19, 2018), http://www.biohealthmatics.com/technologies/hospital-information-systems/clinical-information-systems.

5. For the past and present role of biobanks and the challenges going forward, see Yvonne G. De Souza & John S. Greenspan, *Biobanking Past, Present and Future: Responsibilities and Benefits*, 27 AIDS 303, 303 (2013).

6. Between March 8 and 11, 2018, German doctors debated loosening the ban on remote treatment at the 121st annual meeting of German doctors in Erfurt, Thuringia. Personal treatment is still supposed to be the "gold standard," but they eventually came up with a solution that will allow doctors to treat patients only using electronic communication without seeing them first. *Mediziner beschliessen Ausbau von Onlinesprechstunden*, SPIEGEL ONLINE (May 10, 2018, 8:45 PM), http://www.spiegel.de/gesundheit/diagnose/aerztetag-mediziner-beschliessen-mehr-onlinesprechstunden-a-1207181.html.

7. Charles Auffray et al., *Making Sense of Big Data in Health Research: Towards an EU Action Plan*, 8 GENOME MED. 1, 1 (2016).

factors, and biological-genetic predisposition are developed under the terms "precision medicine"[8] and "pharmacogenetics," *i.e.*, medication tailored to the genetic conditions of the patient.[9]

All this makes the medicine of the Twenty-first century more data-intensive and personal.[10] The collection of health data offers opportunities for better analysis of disease factors, improved diagnostic methods, and higher chances of finding cures as well as a reduction of medical costs and an increased efficiency of the healthcare system.[11] At the same time, patients, physicians, researchers, and developers face new risks as a result of the processing of the collected data.[12] These risks range from discrimination and stigmatization based on the data obtained to unwanted knowledge about one's own health and even the fraudulent misuse of data or blackmailing.[13] Twenty-first century medicine challenges the current legal system.

The handling of health data is regulated via data protection regulations, privacy laws, and medical confidentiality requirements. In recent years, lawmakers in many countries have passed more and stricter data protection and privacy laws.[14] This poses a potential threat to future data-based and preventive medicine. Rigid data privacy and security requirements can render the collection of patient data and its exchange between different healthcare facilities complex, labor-intensive, and costly. This can result in delays in the treatment process. Also, entities are discouraged from sharing personal data, which can result in less informed diagnoses and treatment decisions. Even without strict data privacy laws, healthcare providers feel generally obligated to keep health information confidential and implement data security measures given the potential effects of data security breaches on individual patients.[15] At the same time, the three areas of medicine—

---

8. Francis S. Collins & Harold Varmus, *A New Initiative on Precision Medicine*, 372 N. ENG. J. MED. 793, 793 (2015).

9. Stuart A. Scott, *Personalizing Medicine with Clinical Pharmacogenetics*, 13 GENETIC MED. 987, 987 (2011).

10. Wullianallur Raghupathi & Viju Raghupathi, *Big Data Analytics in Healthcare: Promise and Potential*, 2 HEALTH INFO. SCI. & SYS. 1, 1-2 (2014).

11. Back in 2011, a report of the McKinsey Global Institute estimated the value big data could potentially create in the U.S. healthcare system alone to be over $300 billion per year with more than one third coming from cost reductions. *See* J. MANYIKA ET AL., BIG DATA: THE NEXT FRONTIER FOR INNOVATION, COMPETITION, AND PRODUCTIVITY, MCKINSEY GLOBAL INSTITUTE 2,7 (2011).

12. *See*, *e.g.*, Marco Viceconti et al., *Big Data, Big Knowledge: Big Data for Personalized Healthcare*, 19 IEEE J. OF BIOMED. & HEALTH INFORMATICS 1209 (2015).

13. For an overview of healthcare data breaches in the year 2018, see *The Biggest Healthcare Breaches of 2018 (So Far)*, HEALTHCARE IT NEWS (March 2018), https://www.healthcareitnews.com/projects/biggest-healthcare-data-breaches-2018-so-far.

14. Tim Mullahy, *Privacy Law Is Growing More Extensive – Here's What That Means For Healthcare*, TRIPWIRE (Nov. 27, 2018), https://www.tripwire.com/state-of-security/healthcare/privacy-law-healthcare.

15. Annie Qureshi calls a healthcare data breach "one of the most violating types of data breaches one can encounter." *See* Annie Qureshi, *Healthcare Data Breaches: What Are*

namely, the treatment of diseases, their prevention, and the further development of existing treatment and prevention methods—continue to demand comprehensive data processing.[16]

This article looks for a healthy balance between data access and data protection in the interest of patients' health and privacy. It seeks to identify ways to protect health data privacy without excessively hindering healthcare and medical progress. After this introduction (I), this article examines current approaches to data protection regulation, privacy law, and the protection of patient confidentiality (II), risks associated with the processing of health data (III), needs to protect patient confidence (IV), risks for healthcare and medical progress (V), and possible solutions (VI). It concludes with an outlook and call for healthier approaches to data protection (VII).

II. DATA PROTECTION REGULATION, PRIVACY LAW, AND PATIENT CONFIDENTIALITY

The protection of one's privacy and data is a global issue. Many international treaties and national constitutions refer, expressly or impliedly, to privacy as a core principle or objective.[17] On the national level, the regulations differ from each other, especially when it comes to health data. In Europe, the processing of health data is regulated by general, omnibus data processing regulations. In the United States, sector- and harm-specific privacy laws apply. In most countries, in addition to data privacy and data protection laws, physicians have to protect patient confidentiality under laws or regulations concerning professional responsibilities.

A. *Data Protection Regulation and Privacy Law*

1. Europe

With data protection laws, European countries seek to protect the individual (the "data subject") from potential adverse effects of automated data processing by generally prohibiting the processing of personal data, subject to only enumerated exceptions. The member states of the European Union ("EU") harmonized data protection law with the EU General Data Protection Regulation 2016/679 ("GDPR"), which is directly applicable in all

the Risks?, HEALTHWORKS COLLECTIVE (April 8, 2018), https://www.healthworkscollective.com/healthcare-data-breaches-what-risks.

16. STANFORD MED., *supra* note 3.

17. For an overview, *see* Lothar Determann, *Privacy and Data Protection*, MOSCOW J. INT'L L. 18, 20-21 (2019).

member states of the European Economic Area ("EEA")[18] and applies to companies within and outside the EEA in most cases where either the controller, processor,[19] or data subject is based in the EU.[20]

According to EU data protection law, organizations must not process personal data unless they can provide a justification expressly recognized by law. EU lawmakers reversed the general presumption of liberty (everything is allowed if it is not prohibited)[21] for the field of data processing and data protection law; now, processing of personal data is prohibited unless it is permitted. According to the GDPR, organizations must not process personal data unless they meet all requirements of the regulation and national laws and they can claim one of six "legal bases": (a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes; (b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract; (c) processing is necessary for compliance with a legal obligation to which the controller is subject; (d) processing is necessary in order to protect the vital interests of the data subject or of another natural person; (e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller; (f) processing is 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.[22]

To be valid, consent must be a "freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her."[23] Additional requirements are noted in Article 7 of the GDPR. Most importantly, any respective declaration has to be unambiguous and based on a clear and easily accessible request for consent. Also, the consent must be revocable at any time.

---

18. EU member states plus Iceland, Liechtenstein, and Norway. *See EEA Agreement*, EFTA, https://www.efta.int/eea/eea-agreement (last visited Feb. 29, 2020).

19. These two terms are defined in Article 4 (7) and (8) of the GDPR with the "controller" being a "natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data" and the "processor" is a "natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller." Council Regulation 2016/679, art. 3(1), 4(7)-(8), 2016 O.J. (L 119/32-33).

20. *Id.* art. 3.

21. *See* Jeremy Waldron, *The Rule of Law*, STAN. ENCYCLOPEDIA PHIL. (2016), https://plato.stanford.edu/archives/fall2016/entries/rule-of-la.

22. Council Regulation 2016/679, art. 6(1), 2016 O.J. (L 119/36).

23. *Id.* art. 4(11).

European data protection laws are intended to prevent or restrict the processing of personal data as much as possible, even with regard to publicly available data. "Personal data" includes any information related to an identified or identifiable person, such as a person's name on a business card, a date of birth combined with a patient ID only known to the treating hospital, an image on security footage, or the IP address of a device that belongs to one individual.[24] According to this broad definition, the data itself does not have to identify the data subject – it is enough if it *can* be associated with an identifiable person. The Court of Justice of the European Union has decided that data has to be considered "personal" if the company or body collecting the data in question "has the legal means which enable it to identify the data subject with additional data which the [company or body] has about that person."[25] Data ceases to be "personal" only if it is aggregated in a statistical statement or redacted so thoroughly that it can no longer be connected with an identifiable individual. Effective anonymization is very difficult to achieve in practice,[26] however, and aggregated or anonymized data is far less useful for medical treatment, research, or development than personal data.

"Processing of personal data" is broadly defined as any act related to personal data, such as the collection, storage, transmission, linking or deletion of such data, whether manual or automated, and even redaction and deletion are regulated and restricted as "processing."[27] Any data processing action is subject to extensive restrictions. For example, with regard to the type and frequency of use, authorized users, and storage periods, processing must be limited to what is required for the purpose of processing (principle of data minimization) and must be carried out transparently (principle of transparency) and for a clearly defined purpose (principle of purpose limitation).[28] Data can only be collected for specific and precisely defined purposes while the amount collected has to be as small as possible and the data subject must be granted broad access to the entire process whenever requested.

These requirements are diametrically opposed to the concept of "big data," which relies on large volumes of data that were frequently collected for other purposes and are now analyzed for new insights or connections (e.g.,

---

24. *Id.* art. 4(1).

25. The court qualified an IP address as "personal data" based on the definition used in the Data Protection Directive of the EU/Directive 95/46/EC which corresponds to that of the GDPR. The operator of a website does not have to be able to identify a user himself to qualify the IP address as "personal data." It is enough if public authorities are able to do so with the help of the network provider. *See* Case C-582/14, *Beyer v. Bundessrepublik Deutschland*, 2016 E.C.J. I-779.

26. *See infra* Part II(2).

27. *See* Commission Regulation 2016/679, art. 4(2), 2016 O.J. (L 119) 33.

28. *Id.* art. 5(1)(a-c).

correlations of exposure to a certain part of town to the development of particular diseases after new scientific discoveries create suspicions regarding potential causal links).

Individuals affected by data processing are entitled to extensive rights. They have a right to disclosures and access,[29] meaning they have to be informed whether their data is processed, which data in particular is processed, the purposes of data processing, and the recipients along with other details. In addition, they have a right of rectification, data portability[30] and erasure.[31] The "right to be forgotten" contemplates certain exceptions in Article 17(3)(c) and (d) for public health and scientific research,[32] but an entity that relied on patient consent to collect data in the first place will lose the legal basis for retaining such data if the data subject withdraws her consent, which she is entitled to do at any time.

EU law also imposes far-reaching restrictions on the transfer of data abroad.[33] Cross-border transfers are of significant importance for global medical studies and international research projects and are generally prohibited in Chapter V of the GDPR, subject to complex and narrow exceptions.[34] The EU Commission has only recognized a few countries as providing for an "adequate" level of data protection.[35] Such positive adequacy decisions have been made regarding Andorra, Argentina, Canada, the Faroe Islands, Guernsey, Israel, the Isle of Man, Japan, Jersey, New Zealand, Switzerland, and Uruguay.[36] Regarding the United States, the EU Commission has issued only a limited adequacy decision with respect to companies that register voluntarily under the EU-U.S. Privacy Shield program that the U.S. Department of Commerce administers and whose principles the Federal Trade Commission enforces.[37]

In the absence of an adequacy decision by the EU Commission, organizations can achieve an adequate level of data protection by implementing "appropriate safeguards." Article 46(2) of the GDPR sets out a list of safeguards acceptable to the EU, which include: legally binding and enforceable

29. *Id.* art. 15(1).

30. *Id.* art. 20.

31. *Id.* art. 17.

32. Namely if the processing is necessary for (the vague concept of) "public interest" and "scientific or historical research purposes or statistical purposes." *Id.* art. 17(3).

33. *See* LOTHAR DETERMANN, DETERMANN'S FIELD GUIDE TO DATA PRIVACY LAW 31 (2020).

34. *See* Commission Regulation 2016/679, arts. 44–49, 2016 O.J. (L 119) 60–65.

35. *See id.* art. 45(1).

36. *Adequacy Decisions: How the EU Determines if a Non-EU Country Has an Adequate Level of Data Protection*, EURO. COMMISSION, http://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_de#dataprotectionincountriesoutsidetheeu (last visited Mar. 1, 2020).

37. European Commission, Commission Implementing Decision (EU) 2016/1250 of July 2016, 2016 O.J. (L 207) 11.

instruments between public authorities or bodies, binding corporate rules, standard contractual clauses promulgated by the EU Commission, an approved code of conduct, and approved certification mechanisms.[38] A transfer without an adequacy decision by the Commission or the described safeguards is only possible under certain circumstances listed in Article 49(1) GDPR. Explicit consent of the data subject is one of such exceptions, but the EU Data Protection Board has opined that consent is only appropriate in specific circumstances and not available as a general compliance mechanism.[39]

With respect to health data, which is defined in European law as "personal 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,"[40] the regulation is even more complex and there are further restrictions. Health data is classified as a "special category" as described in Article 9(1) of the GDPR, regardless of how sensitive information pertaining to one's health is; for example, glasses or a Band-Aid visible on security footage or a scanning of a public road by an autonomously-driving car will turn the entire data set into one containing "special categories of personal data," because they contain information on an individual's health. Processing of health data is prohibited unless the data subject has given her explicit consent for one or more specified purposes or the processing is necessary for health or medical purposes, in which case the data may be "processed by or under the responsibility of a professional subject to the obligation of professional secrecy" without consent.[41]

For medical research, Article 9(2)(j) and Article 89 of the GDPR contain several exceptions to the general rules concerning consent and access rights while still requiring certain safeguards. Articles 9(2)(h), 9(2)(j), and 89 of the GDPR allow EU Member States to legislate additional derogations from several provisions of the GDPR concerning the rights of data subjects. For example, in Germany, Article 22 of the new German Federal Data Protection Act (Bundesdatenschutzgesetz) ("BDSG") names several exceptions from the GDPR requirements for processing data of a special category. Such processing is allowed in Germany, for example, for social security administration purposes and for preventive medicine and public health interests, as well as to prevent public harm. At the same time, various necessary safeguards are described in Section 2 of the BDSG, pseudonymization being

38. *See* Lothar Determann et al., *The EU-U.S. Privacy Shield Versus Other EU Data Transfer Compliance Options*, 15 DATA PRIV. & SEC. L. REP. (BNA) (Sept. 5, 2016).

39. *See* European Data Protection Board, *Guidelines 2/2018 on Derogations of Article 49 Under Regulation 2016/679*, GUIDELINES (May 25, 2018), http://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_2_2018_derogations_en.pdf.

40. *See* Council Regulation 2016/679, art. 4(15), 2016 O.J. (L 119/34).

41. *Id.* art. 9.

one of them.[42] Also, Article 27 of the BDSG justifies a limitation on the rights of the data subject under the GDPR if the data processing is necessary for research purposes and the interests of the controller outweigh those of the data subject. However, in addition to the safeguards already mentioned in Section 22 of the BDSG, special categories of personal data have to be anonymized as soon as possible after they have been processed for their original purpose.[43] The leeway granted to EU Member States to legislate derogations from the GDPR enables the creation of a legal patchwork that makes it more difficult for research institutions and companies to conduct international studies or exchange data across borders.

In addition to data protection laws, treating physicians and researchers have to comply with Regulation (EU) 536/2014 on Clinical Trials on Medicinal Products for Human Use, which standardizes authorization procedures, safety necessities, and the requirements for consent to participate in clinical trials.[44] This further complicates the preparation of consent forms and adds restrictions to the subsequent use and sharing of information derived from clinical trials.

2. The United States

In contrast to the situation in Europe, data processing is generally permitted in the United States.[45] There is no comprehensive federal legislation on privacy or data protection comparable to the GDPR. The U.S. Constitution covers and protects certain aspects of privacy in its Fourth Amendment, but without explicitly using the term.[46] Privacy risks and harms are addressed in sector- and harm-specific privacy laws, which are tailored to industries and risks.[47]

Health data is subject to the provisions of the Health Insurance Portability and Accountability Act ("HIPAA"), which requires healthcare sector organizations to implement detailed technical and organizational security measures, disclose data processing practices to patients, and obtain consent in certain limited situations. HIPAA was adopted to "ensure the portability of health benefits when workers change or lose their jobs and will protect workers against discrimination by health plans based on their health sta-

---

42. Bundesdatenschutzgesetz [BDSG][Federal Data Protection Act], June 30, 2017, BGBL I at 2097, amended by Nov. 20, 2019, BGBL I at 1626, art. 22(1) (Ger.). For a definition of "pseudonymization," see Part II(2)(a) below.

43. *Id.* art. 27(3).

44. Council Regulation 536/2014, 2014 O.J. (L 158).

45. *See* Paul M. Schwartz, *Preemption and Privacy*, 118 YALE L.J. 902, 910-16 (2009).

46. *See, e.g.*, Carpenter v. United States, 565 U.S. 400 (2012); Lawrence v. Texas, 539 U.S. 558 (2003); Griswold v. Connecticut, 381 U.S. 479 (1965).

47. *See* DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 257, 349-61 (5th ed. 2014); *see* ANDREW B. SERWIN, INFORMATION SECURITY AND PRIVACY: A GUIDE TO FEDERAL AND STATE LAW AND COMPLIANCE § 7.1, § 28 (2014).

tus."[48] It covers hospitals, healthcare providers, and their billing service providers.[49] Its Privacy Rule regulates the use of protected health information ("PHI")[50] - defined in 45 CFR § 160.103 as any information held by the covered entities regarding health status, provision of health care, or health care payment that can be linked to any individual.[51] The disclosure of such information is based on the patient's express written authorization,[52] with some exceptions for specific circumstances, such as a legal requirement for disclosure.[53] Under HIPAA, patients have rights to access their health data, but, unlike under the GDPR, no "right to be forgotten" or erasure.[54]

In addition to HIPAA, the Federal Policy for the Protection of Human Subjects, also known as the "Common Rule," published in 1991 and revised in 2017 (the revised version went into effect on January 21, 2019), sets standards for "all research involving human subjects conducted, supported, or otherwise subject to regulation by any Federal department or agency."[55] It applies to all research that involves identifiable data about living individuals and is conducted, funded, or regulated by one of the Common Rule departments or agencies.[56] To comply with the Common Rule, a research institution has to submit a written assurance of compliance with the Common Rule to the head of a department or agency[57] and the study has to be reviewed by an Institutional Review Board ("IRB")[58] at least once a year, making sure that the risks for the data subjects are minimized and informed consent has been obtained. In case of non-compliance, the funding of a study[59] or its approval[60] can be terminated. Besides that, the Common Rule does not grant

48. William J. Clinton, *Statement on Signing the Health Insurance Portability and Accountability Act of 1996*, AMERICAN PRESIDENCY PROJECT, www.presidency.ucsb.edu /documents/statement-signing-the-health-insurance-portability-and-accountability-act-1996.

49. The definitions of "covered entities" and "business associates" can be found in 45 C.F.R. § 160.103 (2019).

50. *See Summary of the HIPAA Privacy Rule*, Health Information Privacy, HHS, https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html (last updated July 26, 2013).

51. § 160.103.

52. *Id.* § 164.508(a).

53. *Id.* § 164.502(a)(2).

54. Stacey A. Tovino, *The HIPAA Privacy Rule and the EU GDPR: Illustrative Comparisons*, 47 SETON HALL L. REV. 973, 990 (2017).

55. 45 C.F.R. § 46.101(a) (2019).

56. The definitions of "research" and "human subject" are available in *id.* § 46.102(l) and § 46.102(e)(1) respectively.

57. *Id.* § 46.103(a).

58. *Id.* § 46.107. An IRB has to have least five members of varying backgrounds, at least one scientific and non-scientific member and at least one not affiliated with the institution. *See id.* § 46.107.

59. *Id.* § 46.123.

60. *Id.* § 46.113.

research participants legal remedies in regard to their participation.[61] It should also be noted, that certain low-risk studies conducted by HIPAA-covered entities are exempted from the Common Rule.[62]

Numerous additional laws apply at the state level. For example, the State of California has issued regulations for online services in the form of the Confidentiality of Medical Information Act ("CMIA").[63] It also covers providers of software, hardware, and online services as "providers of healthcare" and provides in California Civil Code 56.10(a) that "a provider of health care, health care service plan, or contractor" is not allowed to "disclose medical information regarding a patient," unless it is authorized by the patient.[64]

On June 28, 2018, California passed the California Consumer Privacy Act ("CCPA") that is set to come into force on January 1, 2020. Already subject to harsh criticism,[65] the CCPA affects most companies worldwide and poses significant organizational challenges.[66] Any business is covered if it does business in California, remotely or with a physical presence, obtains any personal information of any California resident, and exceeds one of three thresholds, (1) generates annual gross revenues of $25 million, (2) processes personal information of 50,000 or more California residents, households, or devices annually, or (3) generates 50% of annual revenue from selling California residents' personal information.[67] "Personal information" is defined more broadly than even "personal data" is defined under the GDPR as "any information that . . . relates to . . .

> a particular consumer or household."[68] "Selling" is any disclosure of personal information or making available personal information for monetary or other valuable consideration. This definition of selling thus covers most types of data exchanges in practice, given that businesses tend to share information only for consideration and consideration is a basic element of any contract.[69] "Consumer"

61. Jessica L Roberts & Valerie Gutmann Koch, *Law vs. Regulations in the Common Rule*, YALE J.L. & TECH. BLOG (Jan. 6, 2016), https://yjolt.org/blog/law-vs-regulations-common-rule.

62. § 46.104(d)(4).

63. CAL. CIV. CODE § 56 (West 2020).

64. *Id.* § 56.10(a).

65. Eric Goldman, *An Introduction to the California Consumer Privacy Act (CCPA)*, in SANTA CLARA UNIV. LEGAL STUDIES RESEARCH PAPER 1 (2018).

66. Lothar Determann, *New California Law Against Data Sharing*, 35 COMPUTER & INTERNET LAW 1, 2 (2018).

67. California Consumer Privacy Act, Cal. Civ. Code § 1798.140(c)(1) (West 2020).

68. *Id.* § 1798.140(o).

69. *Id.* § 1798.140(t). Provides for some exceptions, including consumer-directed disclosures to third parties that do not sell the personal information, limited sharing with service providers, and business transfers in bankruptcy, mergers and acquisitions ("M&A"), and simi-

means any resident, including employees, representatives of businesses, students, or patients.[70] Businesses are required to comply with numerous prescribed requirements, including disclosures using exact statutory language and organization, the implementation of an online opt-out link to enable consumers to prohibit the sale of their personal information,[71] and the granting of new data subject rights, including access, erasure, and portability,[72] as well as channels to exercise the (e.g. a toll-free telephone number). For companies or entities governed by HIPAA or CMIA, the CCPA contains limited exceptions, but they still have to comply with the CCPA to the extent they do not act as a "covered entity" or a "provider of healthcare" (e.g., regarding employees or website visitors) or they process data that does not qualify as "protected health information" under HIPAA or "medical information" under CMIA.[73] Thus, even healthcare providers have to comply with the CCPA. And, more importantly, other businesses are discouraged from sharing data with organizations in the healthcare sector due to the rigid restrictions imposed by the CCPA on data sharing. As a consequence, less information will be available to researchers and developers when the CCPA takes effect.

3. China

In China, the concepts of data protection and data privacy laws are relatively new. During a period of uncompromising growth, data protection regulations and privacy laws were few, fragmented, and sector-specific.[74] Personal data of more than a billion citizens has been available for collection and processing by Chinese researchers and businesses without restrictions similar to those found in Europe or the United States.[75] While lawmakers in the EU and United States are more concerned with individual privacy, regulating data processing, and restricting technology companies,

---

lar transactions. To avail themselves of such exceptions, however, most business will have to adapt their existing contracts and processes.

70. *Id.* § 1798.140(g).

71. *Id.* § 1798.135(a)(1).

72. *Id.* § 1798.100, 105.

73. *Id.* § 1798.145(c).

74. Graham Greenleaf & Scott Livingston, *China's New Cybersecurity Law – Also a Data Privacy Law?*, 144 PRIVACY LAWS & BUS. INT'L REP. 1-7 (2016).

75. Luxia Zhang et al. provide an overview of the increasing role of big data in medical research in China and name respective regulation as a challenge going forward by suggesting to "follow Confucian doctrine to ensure that we obtain true value for medicine - that is, to learn extensively, inquire carefully, think deeply, discriminate clearly, and practice faithfully." *See* Luxia Zhang et al., *Big Data and Medical Research in China*, 360:j5158 BMJ CLINICAL RES. 3 (Feb. 5, 2018).

the Chinese government is encouraging large Chinese technology companies to advance and intensify their data processing to gain leadership positions in artificial intelligence.[76]

Most recently, China has focused on data security and local data availability. In June 2017, the Chinese Cybersecurity Law ("CSL") came into effect, followed by the Personal Information Security Specification (the "PI Security Specification") in May 2018, a standard that will be used as a measure of compliance with China's existing data protection rules.[77] With these measures, the Chinese government seeks to ease access to personal data for the Chinese government, to safeguard social stability, to facilitate censorship and surveillance, and to protect domestic industries from global competition.[78]

The definitional scope of the Chinese Cybersecurity law is comparable to that of the GDPR and CCPA, covering any "personal information," defined as "all kinds of information, recorded electronically or through other means, that taken alone or together with other information, is sufficient to identify a natural person's identity, including but not limited to natural persons' full names, birth dates, national identification numbers, personal biometric information, addresses, telephone numbers, and so forth" in Article 76(5) of the CSL.[79] The information does not have to identify the data subject directly ("or together with other information").[80] The scope is further extended by the PI Security Specification, referring to "all kinds of information . . . to identify a specific natural person or reflect activities of a specific natural person."[81] Health data is not considered as a special category of data and therefore not regulated specifically. Instead it is explicitly named as one example for "personal information" in the PI Security Specification.[82]

Overall, data protection in China is not seen primarily as an instrument to secure the individual's privacy but first and foremost as a matter of and a tool for national security[83] and to protect Chinese data from foreign influ-

---

76. Meng Jing & Sarah Dai, *China Recruits Baidu, Alibaba and Tencent to AI "National Team"*, S. CHINA MORNING POST (Nov. 21, 2017), https://www.scmp.com/tech/china-tech/article/2120913/china-recruits-baidu-alibaba-and-tencent-ai-national-team.

77. Graham Greenleaf & Scott Livingston, *China's Personal Information Standard: The Long March to a Privacy Law*, 150 PRIVACY LAWS & BUS. INT'L REP 25 (2017).

78. For an in-depth discussion of these concerns, see Jyh-An Lee, *Hacking into China's Cybersecurity Law*, 53 WAKE FOREST L. REV. 57 (2018).

79. *See* Zhonghua Renmin Gongheguo Wangluo Anquan Fa (中华人民共和国网络安全法) [Cybersecurity Law of the People's Republic of China], art. 76(5) (2016).

80. *Id.*

81. To which extent the references to "activities of a specific person" actually changes the understanding of "personal information" is debatable, but it can at least in theory be *see*n as an extension of the respective definitions in other regulations. *See* Greenleaf *supra* note 77, at 25-28.

82. *See* Geren Xinxi Chujing Anquan Pingu Banfa (个人信息出境安全评估办法) [China's Personal Information Security Specification], § 3.1 (2017).

83. *See* Cybersecurity Law of the People's Republic of China, art. 1.

ence.[84] According to Article 37 of the Cybersecurity Law, "personal information" and "important data" must be stored within mainland China and can only be transferred outside of China if "the measures jointly formulated by the State cybersecurity and information departments and the relevant departments of the State Council to conduct a security assessment" are followed.[85] The Article refers to "critical information infrastructure" ("CII").[86] Article 31 gives some examples for such CII, including "other critical information infrastructure," and leaves it to the State Council to formulate the specific scope.[87] In a later draft on "Regulations on Protection of Critical Information Infrastructure," healthcare is named as one of the industries considered CII.[88]

In China, privacy and control over one's own data is not perceived as much of an issue as in Europe or the United States.[89] While lawmakers in Europe keep adding restrictions on data processing to protect individual privacy, China is working on implementing a "social credit system" by 2020,[90] which already shows early effects.[91] With this system, everyone's behavior in different areas of life will be evaluated by giving points that add up to an individual score, all by using and processing huge amounts of data.[92]

B. *Consent and Anonymization*

Around the world, healthcare providers and researchers rely primarily on three measures to protect patient privacy: they limit the use of health information to what is necessary to treat the patient, they obtain consent from the patient to secondary or unusual data usages, and they redact or aggregate information so that the individual cannot be identified or associated with the

84. Lee *supra* note 78, at 69.

85. Cybersecurity Law of the People's Republic of China, art. 37.

86. *Id.*

87. *Id.* art. 31.

88. Graham Webster et al., *Critical Information Infrastructure Security Protection Regulations*, CHINA COPYRIGHT & MEDIA (July 12, 2017), https://chinacopyright andmedia.wordpress.com/2017/07/10/critical-information-infrastructure-security-protection-regulations.

89. Hui Zhao & Haoxin Dong, *Research on Personal Privacy Protection of China in the Era of Big Data*, 5 OPEN J. SOC. SCIS. 139, 144 (June 19, 2017).

90. Alexandra Ma, *China Has Started Ranking Citizens with a Creepy "Social Credit" System — Here's What You Can Do Wrong, and the Embarrassing, Demeaning Ways They Can Punish You*, BUS. INSIDER (Oct. 29, 2018, 12:06 PM), https://www.businessinsider.com /china-social-credit-system-punishments-and-rewards-explained-2018-4.

91. Don Reisinger, *China Banned 23 Million People From Traveling Last Year for Poor "Social Credit" Scores*, FORTUNE (Feb. 22, 2019, 9:34 AM), http://fortune.com/2019/02 /22/china-social-credit-travel-ban.

92. Bernard Marr, *Chinese Social Credit Score: Utopian Big Data Bliss Or Black Mirror On Steroids?*, FORBES (Jan. 21, 2019, 12:37 AM), https://www.forbes.com/sites /bernardmarr/2019/01/21/chinese-social-credit-score-utopian-big-data-bliss-or-black-mirror-on-steroids/#b85663a48b83.

data. Patients and their privacy can be protected with one or more of these precautions in combination.[93]

Steps to redact or de-identify personal data are certainly helpful in protecting patient privacy, although the term "anonymization" is usually a euphemism. Seeking consent from a patient can strengthen information self-determination rights and by extension perceptions of privacy. But, the effectiveness of anonymization and consent approaches are increasingly being put to the test by the acquisition and processing of ever-larger amounts of health data in the age of modern medicine.[94]

1. Anonymization

Anonymization means "remov[ing] identifying information from (something, such as computer data) so that the original source cannot be known."[95] The term is frequently used in privacy policies and consent forms, but not usually in statutes. In the GDPR, for example, "anonymization" is not used or defined. Article 4 of the GDPR only defines pseudonymization as "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."[96] Pseudonymization is referred to throughout the GDPR several times as a recommended data security measure. "Anonymous data" is simply data that is not "personal data" for purposes of the GDPR, because it does not relate to an identifiable individual.[97]

Similarly, the CCPA does not use the term "anonymization" but refers to "deidentification" and "aggregation."[98] "Aggregate consumer information" is defined as information that "relates to a group or category of consumers, from which individual consumer identities have been removed, that is not linked or reasonably linkable to any consumer or household, including via a device" and does not mean "one or more individual consumer records that have been deidentified."[99] Aggregate information typically con-

93. *See* Naya Sethi & Graeme T. Laurie, *Delivering Proportionate Governance in the Era of eHealth: Making Linkage and Privacy Work Together*, 13 MED. LAW INT. 168 (2013).

94. Menno Mostert et al., *Big Data in Medical Research and EU Data Protection Law: Challenges to the Consent or Anonymise Approach*, 24 EUR. J. HUM. GENETICS 956 (2016).

95. *Anonymize*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/anonymize, (last visited February 4, 2019).

96. Council Regulation 2016/679, art. 4, 2016 O.J. (L 119/33-35).

97. Even though not defined in the Articles of the GDPR, Recital 26 states that the GDPR is not applicable to "anonymous data." Council Regulation 2016/679, recital 26, 2016 O.J. (L 119/5).

98. *See* California Consumer Privacy Act, CAL. CIV. CODE § 1798.140(o), (h) (West 2020).

99. *Id.* § 1798.140(a).

sists of statistical information concerning a sufficiently large number of persons (e.g., X% of males older than 50 in country Y are overweight) that does not imply information on a particular individual. Aggregate information falls outside the scope of the CCPA and most other privacy laws. The deidentification standard in the CCPA, however, can hardly be met in practice, because any information can be associated with an individual; only statistical statements can be detached from individuals.[100]

When evaluating the actual value and usability of aggregated, redacted, de-identified, pseudonymized, or anonymized data, differences apply with respect to diagnosis and treatment of patients on the one hand, and on the other hand to the use of data for research and development purposes. In the case of medical treatment and care of a specific patient, be it by means of telemedicine using remote communication techniques or the "classic" treatment at a doctor's office, anonymization of health data is often not an option because of the risk of losing the connection to the patient. Every treatment is based on the specific physical and mental condition of the individual. Laboratory values, test results, or X-ray images can only lead to a reliable diagnosis when they are unequivocally connected to a patient. Therefore, complete anonymization would make it nearly impossible to treat a patient if any kind of medical diagnostics should be necessary. Partial redaction of lab reports or images are recommended for data security and privacy purposes (e.g., replacement of patient names with ID numbers), but such measures introduce additional risks for misidentification and never completely rule out the patient's identification. Also, researchers find data sets more useful if they contain as much individual information as possible. If they want to prove a new correlation or causal connection, researchers are interested in information that relates to individuals including that which was not previously recorded for this purpose (because no one had thought of the possible causal connection or correlation). The more data sets are redacted, the less valuable they are for research. On the other hand, for data processing in medical research or internal auditing of health care facilities, where the personal identity of the patient is not necessarily needed, it may be less problematic to redact data sets.

Aside from questions regarding desirability and value of de-identified data, organizations find it extremely difficult or impossible to completely anonymize personal data. So long as information relates to one individual, there is a chance that this individual can be identified. Only aggregate information relating to groups can be truly anonymous. Encryption or coding measures such as the already mentioned pseudonymization, are not suffi-

100. *See id.* § 1798.140(h). "Deidentified" means "information that cannot reasonably identify, relate to, describe, be capable of being associated with, or be linked, directly or indirectly, to a particular consumer." Any personal information, however, can be associated with a person.

cient in this regard as long as one person still has the "key" to re-identifying the individual data subject. Erasing all personal references is difficult to achieve due to numerous reference data sources.[101] The data subject might be re-identified using these. The re-identification of the medical data of Massachusetts Governor William Weld in 1997 is a well-known example in this regard.[102] He had collapsed during a public event and a researcher gained access to his medical records from a Massachusetts Group Insurance Commission ("GIC") database by using his zip code and birth date.[103] Essentially, what was considered to be anonymization for a long time is actually de-identification.[104] Some see this concept as insufficient and limited due to the mentioned possibility of re-identification.[105] Others point out that the risk of such re-identification is often so low that de-identification is still a useful tool for data protection.[106] In any case, complete anonymization is hard to achieve and often more of a "myth" than reality.[107] The substantial and continuous increase in available computing power and the easier and faster detection of information from digital sources have facilitated the linking of different data sets and thus enabled identification by merging different data sources.[108]

Also, as advances have been made in human genome research, health data increasingly includes genetic information, which cannot be irrevocably and completely de-identified given the uniqueness of DNA.[109] Several individual studies have shown that redacted genetic information of study participants can still be used to identify individuals with other publicly accessible studies and non-health data, such as the year of birth or place of residence of

---

101. Ira S. Rubinstein & Woodrow Hartzog, *Anonymization and Risk*, 91 WASH. L. REV. 703, 710-11 (2016).

102. For a summary of the incident and critical view on the actual impact of the case on the anonymization debate, *see* Daniel C. Barth-Jones, *The "Re-Identification" of Governor William Weld's Medical Information: A Critical Re-Examination of Health Data Identification Risks and Privacy Protections, Then and Now* (2012) (unpublished manuscript) (available at SSRN: https://ssrn.com/abstract=2076397).

103. *Id.*

104. The National Institute of Standards and Technology ("NIST") describes de-identification as "a tool that organizations can use to remove personal information from data that they collect, use, archive, and share with other organizations." *See* Simson L. Garfinkel, *De-Identification of Personal Information*, NISTIR 8053, at 4 (2015), https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf.

105. *See* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 U.C.L.A. L. REV. 1701 (2010).

106. Jane Yakowitz, *Tragedy of the Data Commons*, 25 HARV. J.L. & TECH. 1, 35-36 (2011).

107. Rubinstein & Hartzog, *supra,* note 101, at 704.

108. *See* Ohm, *supra* note 105.

109. Benjamin E. Berkman et al., *The Ethics of Large-Scale Genomic Research*, *in* ETHICAL REASONING IN BIG DATA: AN EXPLANATORY ANALYSIS 53, 54 (Jeff Collmann & Sorin Adam Matei eds., 2016) (noting that "all genomic data is theoretically identifiable").

the person affected.[110] For example, in 2013, a researcher was able to re-identify more than 40% of a sample of anonymous participants of a DNA study by only primarily using their zip code, gender, and date of birth.[111] Such identification is likely to increase even more with further collection of health data, especially via social media or fitness trackers.

In addition, anonymizing data has several downsides. Much medical research is based on accumulating certain information and then drawing conclusions from it. Statistics can help to detect a possible correlation Z between factor X and outcome Y. But, without a link between X and Y, there is no Z. The individual person is that link. This connection is important in the field of genetic research and public health (e.g., the prevention of epidemics or the fight against widespread diseases such as cardiovascular diseases or diabetes mellitus) in order to gain a better understanding of the respective causes and possible health consequences.[112] Anonymization makes such data worthless for research purposes by separating the data from the person, or at least reduces its value.[113]

Anonymization or pseudonymization is also of limited use in the context of personalized medicine. Medicine is shifting towards a more personal approach, tailoring the treatment to the individual condition and needs of the patient, instead of simply using generalized diagnostic and treatment methods.[114] In this regard, the volume of collected data, the new and repeated linking of data sets,[115] and the amount of detail is crucial.[116] Both are substantially harder to execute or not possible at all when data is anonymized.[117]

---

110. Melissa Gymrek et al., *Identifying Personal Genomes by Surname Inference*, 339 SCIENCE 321, 321 (2013).

111. Adam Tanner, *Harvard Professor Re-Identifies Anonymous Volunteers In DNA Study*, FORBES (February 4, 2019, 8:30 PM), https://www.forbes.com/sites/adamtanner/2013/04/25/harvard-professor-re-identifies-anonymous-volunteers-in-dna-study/#4c4fc70892c9.

112. *See* Sophie Stalla-Bourdillon & Alison Knight, *Anonymous Data v. Personal Data - False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data*, 34 WIS. INT'L L.J. 284, 284 (2016) (noting that "the value or knowledge that can be gained from analyzing datasets (particularly using automatic algorithmic software) is maximized by virtue of finding patterns, basically linking relationships between data points.").

113. *See* Mostert et al., *supra* note 94; Adrian Thorogood et al., *An Implementation Framework for the Feedback of Individual Research Results and Incidental Findings in Research*, 15 BMC MED. ETHICS 1 (2014), http://www.biomedcentral.com/1472-6939/15/88; *see also* Ohm, *supra* note 105, at 1704 (noting that "data can be either useful or perfectly anonymous but never both").

114. *See* Jacob S. Sherkow & Jorge L. Contreras, *Intellectual Property, Surrogate Licensing, and Precision Medicine*, 7 IP THEORY 1 (2018).

115. Brent D. Mittelstadt & Luciano Floridi, *The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*, 22 SCI. & ENGINEERING ETHICS, 303, 314 (2015) (noting that "big data is intended by design to reveal unforeseen connections between data points").

116. In 2017, John Bell, Regius Professor of Medicine at Oxford University, wrote in the U.K. government's life sciences industrial strategy: "One of the most important resources held by the UK health system is the data generated by the 65 million people within it." JOHN

Patients can also be harmed by anonymization. If the data on which research results are based can no longer be connected to a specific person, this person can neither authorize later research projects that are in her own interest, nor be informed about new insights that may affect her and that sometimes require urgent clinical action (e.g., discoveries regarding a previously unknown and now treatable health condition in an unidentifiable study participant).[118] The patient has an elementary interest in such direct participation in the information gained.[119] In addition, knowledge gained retrospectively, which potentially influences the risk or the individual course of a disease, can no longer be retransmitted, causing an ethical dilemma.

The downsides of anonymization apply, to a lesser extent, to pseudonymization as it uses similar techniques and the goal is to at least somewhat loosen the connection between data and data subject. Pseudonymization should be considered as a data security measure but does not completely avoid restrictions under data protection or data privacy law. It also creates risks of its own, including errors in re-identifying patients by numbers on lab reports.

2. Consent

Where healthcare providers and researchers seek consent from patients, they pay respect to patient self-determination and autonomy but face practical implementation problems. Issues concerning individual consent have historically been primarily relevant with respect to medical research and clinical trials, while in the field of treatment, healthcare providers did not need to obtain consent from patients with respect to the limited data processing that was necessary for treatment. But now, after significant medical advances, healthcare providers and patients face additional treatment and preventive maintenance options that require significantly more data processing and increased concerns relating to the necessity and sufficiency of patient consent. This has caused a shift in the discussion towards everyday medical treatment and care.

The development and use of personalized medicine is based on and benefits from the processing and exchange of large amounts of health da-

---

BELL, LIFE SCIENCES: INDUSTRIAL STRATEGY 56 (2017), www.gov.uk/government/publications/life-sciences-industrial-strategy.

117. *See* W. Nicholson Price II, *Big Data, Patents, and the Future of Medicine*, 37 CARDOZO L. REV. 1401 (2016).

118. G. Owen Schaefer & Julian Savulescu, *The Right to Know: A Revised Standard for Reporting Incidental Findings*, 48 HASTINGS CTR. REP, 22, 22 (2018).

119. *Id.*; *see also* Effy Vayena & Alessandro Blasimme, *Biomedical Big Data: New Models of Control Over Access, Use and Governance*, 14 J. BIOETHICAL INQUIRY 501 (2017).

ta.[120] This relates to a wide range of applications and ranges from data processing for the purpose of self-optimization (monitoring of fitness data by the patient) or the supervision of an initiated medical treatment (monitoring of medication administered by the doctor) to a continuous preventative analysis of health data (e.g., observation of vital data to enable early intervention). In the case of chronic diseases, these can be combined with data on the patient's lifestyle, following the idea of an automated diagnosis or even therapy recommendations by artificial intelligence in the future.

Physicians face the problem of obtaining the necessary consent of the data subject, due to the unpredictability and complexity of such ambitious projects. Besides other challenges,[121] they need to decide how extensive and broad the consent should be. They have the option to use declarations of consent that are broadly formulated with regard to objectives and possible uses. These broad types of consent assign the supervision of later renewed or further use of the data to certain bodies, such as in the form of institutional review boards (United States) or research ethics committees (Europe)[122] in medical research. Informing the patients about the objectives and methods of data processing in a detailed way before they grant consent would be another option.

If organizations use open-ended privacy notices and excessively broad scope definitions in consent declaration forms, trying to anticipate potential future data uses, they undermine the validity of the resulting consents. Consent according to data protection regulations has to be a voluntary, unequivocally expressed declaration stating specific purposes. If consent is too broad, it may not meet legal standards. For example, Articles 4(11), 6(1)(a), and 7 of the GDPR set high standards in this regard and in principle demand consent to be narrow and specific even though Recital 33 of the GDPR acknowledges that this might not be possible in medical research and a broader form of consent may be required, creating uncertainty in practice.[123]

Even if an organization provides excessive detail in privacy notices, this by no means guarantees that the consenting patient is truly informed. Patients who are overwhelmed with details they cannot understand due to a lack of medical or genetic expertise may not even try to understand any of

120. Akram Alyass et al., *From Big Data Analysis to Personalized Medicine For All: Challenges and Opportunities*, BMC MED. GENOMICS (June 27, 2015), https://bmcmedgenomics.biomedcentral.com/track/pdf/10.1186/s12920-015-0108-y.

121. Christine Grady describes cultural differences as an example of the potential issues arising from obtaining consent in the modern age of medicine. She refers to the different ways of making a decision, alone or within a family or community, and distinct moral values as reasons for potential misunderstandings. *See* Christine Grady, *Enduring and Emerging Challenges of Informed Consent*, 372 N. ENG. J. MED.855 (2015).

122. Jane Kaye, *The Tension Between Data Sharing and the Protection of Privacy in Genomics Research*, 13 ANN. REV. GENOMICS & HUM. GENETICS 415, 422 (2012).

123. David Townend, *Conclusion: Harmonization in Genomic and Health Data Sharing for Research: An Impossible Dream?*, 137 HUM. GENETICS 657, 661 (2018).

the details in the privacy notice and thus end up less informed than they could have been based on a shorter, broader notice.[124]

Moreover, if an organization tries to create very specific scope definitions in consent declaration forms, this may be counter-productive in the treatment context. Many treatments consist of various separate constituent parts, and could thus each require a separate, specific consent. Patients, who usually assume every diagnostic measure is part of the treatment as a whole and will therefore be carried out in their best interest, may become concerned or irritated if they are repeatedly prompted with consent requests and they may lose trust in the treating physician.

The language barrier is another challenge that commonly arises in the medical field.[125] A growing number of patients are not native speakers or do not speak the language of the country in which they receive treatment.[126] Also, few doctors are specifically trained to obtain informed consent from patients and usually do not have the time to inform them sufficiently about all the details of their treatment and the exact ways their data will be used.[127] In reality, the patient is usually handed a standardized consent form that describes the upcoming procedure and possible risks and asked to sign it. This might be done by a medical student or a nurse and it is not guaranteed that a doctor will be present for a patient to ask any specific questions they may have.[128]

Whether consent is an absolute must for information self-determination has long been questioned. The GDPR, for example, provides many exceptions to consent requirements[129] and the CCPA requires consent only from minors and parents of children.[130]

In the treatment context in particular, it often seems doubtful whether patients declare consent exercising free will.[131] In some situations, there

---

124. *See* Matthew E. Falagas et al., *Informed Consent: How Much and What do Patients Understand?*, 198 AM. J. SURGERY 420, 421 (2009).

125. For an early perspective on the situation in the United States, *See* Glenn Flores, *Language Barriers to Health Care in the United States*, 355 N. ENG. J. MED. 229, 230 (2006).

126. Renata F. I. Meuter et al., *Overcoming Language Barriers in Healthcare: A Protocol for Investigating Safe and Effective Communication when Patients or Clinicians Use a Second Language*, 15 BMC HEALTH SERV. RES. 371 (2015); Allison Squires, *Strategies for Overcoming Language Barriers in Healthcare*, 49 NURSING MGMT. 20, 21 (2018).

127. Grady, *supra* note 121, at 857.

128. *See* Michael Billig, *Medizinstudenten im PJ - Das hätte auch schiefgehen können*, SPIEGEL ONLINE (Feb. 13, 2019), http://www.spiegel.de/lebenundlernen/uni/medizinstudenten-im-pj-das-haette-auch-schiefgehen-koennen-a-914791.html.

129. Winfried Veil gives a good overview over the exceptions regulated in the GDP. *See* Winfried Veil, *Public Interest in the GPDR 2* (Feb. 5, 2019, 6:55PM), https://.flickr.com/photos/winfried-veil/39501609474/in/dateposted-public.

130. California Consumer Privacy Act, CAL. CIV. CODE § 1798.120(c) (West 2020).

131. Ulrich M. Gassner accurately describes "take it or leave it" situations as an example of consent that might not be based on free will entirely. *See* Ulrich M. Gassner, *Informed Consent und Digital Health*, *in* BIG DATA UND E-HEALTH 35 (2017).

might not be an option other than to agree to someone using the data.[132] For example, a patient will regularly have to share personal information throughout treatments, for example, on the development of pain, symptoms, or side effects. Without information sharing, the treatment is not effective or possible. Many patients have no actual choices whether to give consent.[133] Even where choices exist in principle, the doctor is in a stronger position because of her superior knowledge, so the patient will usually respect her authority and follow a request for consent.[134]

Even if one assumes a patient can freely give informed consent, it is debatable whether individual consent is the best basis for determining whether data should or should not be used. By primarily focusing on the seemingly compulsory need for consent whenever data is collected, we lose sight of the fact that information is constantly exchanged as part of our everyday life and we do rely on it being collected and processed.[135] No one owns data,[136] and most people accept a "tension between autonomy and solidarity" with respect to data collection and usage.[137]

According to many data protection laws, organizations have to obtain consent if they want to use data for different purposes than those for which they originally collected the data. The GDPR prescribes in Article 5(1) that "personal data shall be . . . collected for specified, explicit and legitimate purposes and not further processed in a matter that is incompatible with those initial purposes" provided that "scientific . . . or historical research . . . or statistical purposes shall not be considered to be incompatible with the initial purposes."[138] This exception is helpful for scientific, statistical, and historic research, but does not expressly extend to medical research or product development. Also organizations that try to adhere to the purpose limitation principle often impose limits on themselves in privacy policies when they define their original data collection objectives. Such self-imposed purpose limitations often exceed what is expressly required by law. Moreover, researchers and developers are not exempt from the requirement of data minimization.[139]

---

132. Winfried Veil lists applying for a loan as an example. *See* Winfried Veil, *The Emperor's New Clothes - On the Structural Shortcomings of Both the Old and the New Data Protection Law*, 10 NEUE ZEITSCHRIFT FÜR VERWALTUNGSRECHT 686-696 (2018).

133. Charlotte Tschider, *The Consent Myth: Improving Choice for Patients of the Future*, 96 WASH. U. L. REV. 1505, 1506, (2019).

134. Robert Harvey, *"Informing Consent": Challenging Perceptions within Medical Law. Can We Ever Truly "Consent" Under the Present "Law"?* (2016) (unpublished manuscript) (available at SSRN: https://ssrn.com/abstract=2813239).

135. *See generally* Veil, *supra* note 132, at 686.

136. *See generally* Lothar Determann, *No One Owns Data*, 70 HASTINGS L.J. 1 (2018).

137. Townend, *supra* note 123, at 662.

138. Regulation (EU) 2016/679 of May 25, 2018 (General Data Protection Regulation) art. 5(1) (EU).

139. *See id.* art. 89.

The vague and somewhat contradictory wording of the GDPR[140] creates uncertainty and doubt and makes it difficult for organizations to determine how they can comply.[141] To be on the safe side, many organizations consider seeking consent from the data subject before they use existing data for new purposes. But, in many cases, particularly relating to pseudonymized data, the data subject or the surviving relatives are difficult to find or identify. Even in the case of living data subjects whose contact details are known, there is a risk of "fatigue" and a corresponding lack of willingness to respond.[142]

Obtaining consent might sometimes be problematic in the first place, but as long as the respective treatment is still needed or ongoing, there is usually a better chance that the patient will cooperate in additional studies or research of any kind. Once a patient has "left the system," it is more difficult to reach and convince a person to read a lengthy privacy notice and declare consent.[143] A former patient may not want to be reminded of a cured or chronic disease and prior treatment. Patients who were cured based on information of prior patients may be disinclined to "give back" and allow their information to be used for further studies.[144]

C. *Medical Confidentiality*

In addition to data protection regulations, physicians are also obligated to maintain medical confidentiality. Patients expect confidentiality as a "core value of medicine,"[145] which is codified in rules of conduct of the medical profession.[146] According to the Hippocratic oath[147] and the more

140. At what point is research no longer compatible with the initial purpose? How can data minimization and the large amount of information needed for research be reconciled?

141. *See generally* Townend, *supra* note 123.

142. *See* Thomas Ploug & Søren Holm, *Meta Consent: A Flexible and Autonomous Way of Obtaining Informed Consent for Secondary Research*, 350 BMJ 1 (2015).

143. *See* Mostert et al., *supra* note 94, at 957.

144. *See generally* Kaye, *supra* note 122 (discussing concerns relating to patient consent).

145. Thomas H. McCoy & Michael C. Hughes, *Preserving Patient Confidentiality as Data Grow: Implications of the Ability to Reidentify Physical Activity Data*, JAMA NETWORK OPEN (Dec. 21, 2018), https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2719121.

146. *See, e.g.*, *Berufsordnung der Landesärztekammer*, LANDESÄRZTEKAMMER (Mar. 26, 2019), https://www.laekh.de/images/Aerzte/Rund_ums_Recht/Rechtsquellen/berufsordnung.pdf.

147. *See The Hippocratic Oath and Others*, MCMASTER UNI., https://hslmcmaster.libguides.com/c.php?g=306726&p=2044095 (last updated Apr. 26, 2019) ("What I may *see* or hear in the course of the treatment or even outside of the treatment in regard to the life of men, which on no account one must spread abroad, I will keep to myself, holding such things shameful to be spoken about.").

modern Declaration of Geneva,[148] physicians are obligated to maintain secrecy with regard to knowledge acquired in the course of providing healthcare. Moreover, physicians are prohibited by criminal law from releasing personal data in violation of medical profession regulations.[149]

These regulations have two purposes. One, like data protection regulations, medical confidentiality regulations seek to protect the interest of the individual in the confidentiality of certain (health-relevant) facts as an expression of the right to informational self-determination. Two, medical confidentiality is necessary as a basis for general trust in the doctor-patient relationship in order to protect the functioning of the health system. For example, if patients do not trust doctors, they may not seek treatment for contagious diseases and expose themselves and others to unnecessary risks. Also, patients who do not trust their doctor may withhold information or lie about lifestyle habits, for example, downplaying alcohol consumption, denying tobacco use, or exaggerating daily exercise. Based on such misinformation, physicians can be less effective in treating the individual patient and drawing conclusions for general healthcare and lifestyle recommendations.

III. RISKS OF DATA PROCESSING FOR PATIENTS, RESEARCHERS, AND DOCTORS

Risks of data processing have been well researched: Individuals need protection from psychological, economic, and other privacy harms that states, businesses, criminals, and others cause. For example, these harms can be caused by identity theft; blackmail; bullying; stalking; revelation of secret location or identities of spies, domestic abuse victims, or persons in witness protection programs; stigmatization based on addictions, diseases, political opinions, religion, race, or sexual preferences; computer hacking; irritating direct marketing methods; unfair business practices based on surreptitious data collection; and discrimination by employers, banks, and insurance companies based on information about pre-existing health conditions.[150] The main reason for the increasingly extensive data processing

148. *See WMA Declaration of Geneva*, WORLD MEDICAL ASSOCIATION, https://www.wma.net/policies-post/wma-declaration-of-geneva (last updated Oct. 2018) ("I will respect the secrets that are confided in me, even after the patient has died").

149. STRAFGESETZBUCH [STGB] [CRIMINAL CODE] § 203(1) (2019) (Ger.).

150. Danielle Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1915 (2019); *see also* Ryan Calo, *Privacy Harm Exceptionalism*, 12 COLO. TECH. L.J. 361 (2018); Amit Datta et al., *Automated Experiments on Ad Privacy Settings*, PROC. ON PRIVACY ENHANCING TECHS. 92, 92 (2015); Margaret Hu, *Big Data Blacklisting*, 67 FLA. L. REV. 1735, 1809 (2015); Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J.L. & TECH. 148, 151 (2016); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process For Automated Predictions*, 89 WASH. L. REV. 1, 15 (2014); Daniel J. Solove, *Conceptualizing Priva-*

regulation in the field of healthcare and medical research is the high sensitivity of the data. An individual's health is one of the most private aspects of life and the corresponding information is accordingly vulnerable. The collection, use, and sharing of health data creates many risks for patients, researchers, and physicians. In the past, collected health data could be de-identified in a staged order to protect those affected so that their identity was not revealed in the event of unauthorized access. This possibility is subject to increasing uncertainty due to technical developments.[151]

The possible exposure of such data can lead to a severe level of stigmatization, especially in the case of infectious diseases such as HIV[152] or mental health conditions such as depression or schizophrenia.[153] Patients may experience embarrassment, shame, and even social exclusion should information of this nature become public. In contrast to leaked credit card details or a hacked online e-commerce account, the harm resulting from disclosures that a person suffers from a certain disease cannot be undone by changing a password or blocking access to a bank account. Instead, the perceived stigmatization is likely to affect the quality of life of the affected and can also cause additional health conditions, such as a variety of psychosomatic symptoms.[154]

Inadvertent disclosures are not uncommon, yet they are still potentially very harmful.[155] In 2001, a manufacturer of a well-known anti-depression drug ran an email notification service for patients to remind them to take and reorder the drug. In an email to all subscribers, a company employee accidentally inserted the email addresses of all subscribers in the visible "to" field (as opposed to the suppressed "bcc" field), thus introducing all subscribers to each other and revealing their subscription.[156] The FTC launched a complaint and compelled the company to introduce additional security protocols for sensitive data.[157]

In addition to fears regarding data security risks, patients are also concerned that the availability and disclosure of their health information creates

---

cy, 90 CAL. L. REV. 1087 (2002); Daniel J. Solove & Danielle Keats Citron, *Risk and Anxiety: A Theory of Data-Breach Harms*, 96 TEX. L. REV. 737, 738-39 (2018);

151. *See supra* Part II(2)(a).

152. Only recently, the names, addresses, and HIV status of 14,200 people in Singapore were leaked. *Fury at HIV Data Leak in Conservative Singapore*, MED. EXPRESS (FEB. 10, 2019), https://medicalxpress.com/news/2019-02-fury-hiv-leak-singapore.html.

153. Even diseases like cancer that are necessarily associated with a certain (negative) public perception, can cause the feeling of being stigmatized. *See* J. Ernst et al., *Perceived Stigmatization and Its Impact on Quality of Life - Results from a Large Register-Based Study Including Breast, Colon, Prostate and Lung Cancer Patients*, 17 BMC CANCER 741 (2017).

154. Michael Koller et al., *Symptom Reporting in Cancer Patients: The Role of Negative Affect and Experienced Social Stigma*, 77 CANCER 983, 994 (1996).

155. *See, e.g.*, Gavin Yamey, *Eli Lilly Violates Patients' Privacy*, 323 BMJ 65 (2001).

156. Complaint, Eli Lilly & Co., C-4047, at ¶ 6 (Fed. Trade Comm'n May 8, 2002).

157. Decision and Order, Eli Lilly & Co., C-4047, § II (Fed. Trade Comm'n May 8, 2002).

discrimination risks.[158] On the basis of collected health data, insurance companies could burden individual customers with higher rates for health, life, or disability insurance or not accept them at all.[159] Employers could use health information as an opportunity to assess the performance of their employees or to refrain from hiring, retaining, or promoting job candidates. Banks could grant loans only to the healthy or vary the interest depending on the health record of a customer as one without serious diseases is more likely to work longer and therefore to be able to meet his or her contractual obligations. The same applies to housing where a healthy tenant could be seen as the more reliable one. Health information might become an even more important economic factor if it is widely available to businesses.

Another phenomenon is that some individuals prefer not to be confronted with unwanted knowledge about their own state of health by uncovering genetic risks or existing illnesses that can turn their life plans upside down.[160] While the "transparent patient" is desirable from a diagnostic point of view, some could appreciate the freedom of not knowing every detail about their health situation. The collection of an extensive amount of data increases the chance of a random discovery of certain predispositions or an actual illness. Especially when there is no therapeutic consequence, the benefit is at least doubtful and might be outweighed by the potential harm. An increasing number of companies provide basic genetic research to consumers and give easy access to DNA tests. Information that used to be difficult to obtain is suddenly widely available and also concerns family members of the individual who chooses to obtain or publish her own DNA information. Therefore, there is an increasing possibility of unintentionally discovering details about one's genetic predisposition for diseases or regarding family ties, e.g., being adopted or having a different father than expected.[161]

The need to store health information for data processing also makes this data vulnerable to criminal activity. Unauthorized persons can gain access to data stored by health insurance companies, doctors, or research institutions and thus "capture" these data and use them to the disadvantage of the data subject. Patients could be blackmailed with the knowledge of a stigmatized disease.[162] Also, knowledge of health information in the hands of unau-

158. Berkman et al., *supra* note 109, at 59; *see* Ribhi Hazin et al., *Ethical, Legal, and Social Implications of Incorporating Genomic Information into Electronic Health Records*, 15 GENETICS IN MED. 810, 810-15 (2013).

159. Hazin et al., *supra* note 158, at 814.

160. Ellen Wright Clayton, *Incidental Findings in Genetics Research Using Archived DNA*, 36 J.L., MED. & ETHICS 286 (2008).

161. Elle Hunt, *Your Father's Not Your Father*, GUARDIAN (Sept. 18, 2018), https://www.theguardian.com/lifeandstyle/2018/sep/18/your-fathers-not-your-father-when-dna-tests-reveal-more-than-you-bargained-for.

162. Emily Yahr, *Charlie Sheen Says He's HIV-Positive*, WASH. POST (Nov. 17, 2015), https://www.washingtonpost.com/news/arts-and-entertainment/wp/2015/11/17/charlie-sheen-i-am-in-fact-hiv-positive.

thorized persons can be the basis for identity theft or fraudulent offers such as alleged miracle cures or simply counterfeit drugs that try to capitalize on the increased vulnerability of a seriously ill person.

More than just patients face risks associated with processing of data. Treating physicians can suffer reputational damage from public access to "bad" statistics as well as researchers from "false" research results. For example, the provision of data on evaluation portals on the Internet or the statistical processing of treatment information, detached from the treatment context, can lead to an incorrect public perception.[163] This is obvious in the case of intended damage to reputation, but it is also conceivable in the reproduction of information that is accurate at the data level. For example, a high mortality rate among a physician's patients could be read as an indication of an accumulation of treatment errors, but in reality be caused by a specialization or an above average willingness to accept difficult cases of severe diseases associated with increased mortality.[164] Such "false" transparency can cause significant negative consequences for the people involved. Even today, it leads to some institutions rejecting "high-risk patients" because they "endanger the statistics."[165] This puts patients at a concconcrete risk.

The same is true for doctors who are evaluated by their patients, be it online or via internal surveys carried out by a hospital. This can lead to medical treatments being influenced by the patient's wishes and demands, reasonable or not, only to improve evaluation scores. Some physicians might even consider influencing the ratings by making patients sign waivers to prevent them from giving unfavorable reviews and thereby counteract the whole idea of evaluation.[166]

IV. PROTECTING TRUST AND HEALTH DATA

A patient regularly has to share sensitive information about her state of health and lifestyle during a medical treatment or in the context of medical research.[167] The patient's or study participant's willingness to accept the risks inherent in the processing of health data is crucial to provide promising

163. Jennwood Chen et al. describe the discrepancy between reviews given on online platforms and the actual satisfaction of the patient. *See* Jennwood Chen et al., *Online Physician Review Websites Poorly Correlate to a Validated Metric of Patient Satisfaction*, 227 J. SURGICAL RES. 1, 4-5 (2018).

164. *See* Richard Lilford & Peter Pronovost, *Using Hospital Mortality Rates to Judge Hospital Performance: A Bad Idea That Just Won't Go Away*, 340 BMJ 955-57 (2010).

165. *See id.*

166. Julie B. Samora et al., *Physician-Rating Web Sites: Ethical Implications*, 41 J. HAND SURGERY AM. 104, 104 (2016).

167. Angeliki Kerasidou, *Trust Me, I'm a Researcher!: The Role of Trust in Biomedical Research*, 20 MED. HEALTH CARE & PHILOS. 43-50 (2017).

treatment and carry out successful research.[168] However, it also requires trust in the actors of the health system, such as doctors, hospitals, health insurance funds, and research institutions.[169] Protecting this trust is a common fundamental objective of the rules of data protection law and the rules of patient confidentiality.

The patient relies on the confidentiality, security, and accuracy of the information disclosed when health data is processed. However, the concept of trust is not one-sided. The physician also trusts in the accuracy of the data and its secure use, particularly in the context of data processing for research projects that often involve considerable development efforts.[170]

Patients develop two kinds of trust: trust with regard to certain actions ("act-trustworthiness") and trust with regard to certain persons ("character-trustworthiness").[171] In the first case, the patient trusts that a particular action to be taken suits the interests of the patient and the actor. Trust exists as long as the interests of the person concerned and the person acting are perceived to be aligned. In the second case, the patient trusts an individual physician, researcher, or organization because of an impression or reputation of appropriate attitude and accountability.

This distinction may explain the different perception of trust in cases of medical treatment (the patient seeks treatment from a resident physician on the basis of pain or another symptom) on the one hand and on the other hand medical research (a cancer patient participates in a clinical trial to develop a new drug). In the interaction with the treating physician, the patient can usually assess the extent of and reason for collecting his data. This reason is obvious to the patient as she is the one suffering from a certain condition and therefore seeks help in the form of a sufficient treatment. Knowing the existing professional ethical obligation to act in the interest of the patient and the professional and legal obligation to maintain confidentiality, the pa-

168. At least from the (subjective) perspective of the patient, the treatment outcome is better when the level of trust is higher. "Across diverse clinical settings, patients reported to be more satisfied with treatment, to show more beneficial health behaviours, less symptoms and higher quality of life when they had higher trust in their health care professional." Johanna Birkhäuer et al., *Trust in the Health Care Professional and Health Outcome: A Meta-Analysis*, 12 PLOS ONE 1 (2017).

169. Anna C. Mastroianni states, "Maintaining public trust is absolutely crucial to the research enterprise. Without trust, volunteers will be impossible to recruit, and the public will be unwilling to fund research." Anna C. Mastroianni, *Sustaining Public Trust: Falling Short in the Protection of Human Research Participants*, 38 HASTINGS CTR. REP. 8, 8 (2008).

170. Susan Dorr Goold sees trust as "essential to both physician and patient." Susan Dorr Goold, *Trust, Distrust and Trustworthiness: Lessons from the Field*, 17 J. GEN. INTERNAL MED. 79, 79 (2002).

171. J. Patrick Woolley, *Trust and Justice in Big Data Analytics: Bringing the Philosophical Literature on Trust to Bear on the Ethics of Consent*, 32 PHILOS. & TECH. 111 (2017).

tient can assume that the data will be used in his or her own interest.[172] Trust arises both in the action itself and in the person acting.

The situation is different for the patient in a clinical study. Due to a lack of insight and expertise, it is not possible for the patient to be aware of the actions carried out in the context of data processing. Also, the cause of the data collection and its processing is not as present and tangible as it might be in the case of the treatment of an actual disease. This is further intensified through complex and international "big data" applications.[173] The patient cannot determine whether the mutual interests of all parties do coincide without insight into the individual actions. Trust in data processing can therefore only arise with regard to the acting person.[174] Lacking this trust, the patient may get the impression that, for example, due to economic motivation, data processing is not carried out in the patient's interest or is neglected contrary to the patient's interests. This poses a severe threat to a study and its successful outcome,[175] as individuals are less likely to participate. If, in the case of commercial research, the patient gets the impression that her own data is being processed for the financial benefit of another, a feeling of exploitation may arise. These considerations and the establishment of trust in the data processing authority have to be taken into account when solutions are being developed. After all, the issue of trust in the context of a research project is not only relevant to the patient. Moreover, the researching physician usually has no access to the complete data set and must trust in a correct process of data processing. He lacks the ability to control several risks of the research project. Therefore, the physician has to trust in the functioning of individual institutions (such as security committees and data protection officers) and the independence of their actions from the sponsor's economic interests.

Overall, it is important to seek the benefits of efficient and comprehensive data processing without undermining patient confidence, be it within the personal interaction between patient and doctor or during a medical study. Extensive data protection laws can be helpful in that regard, but there is no guarantee they will actually help to build the patient's trust.[176] On the contrary, overly detailed privacy notices and repeated requests for specific consent can even increase levels of distrust. Instead of increasing the frequency of consent requests and the detail in privacy notices, the focus should be put on ensuring the trustworthiness of doctors and research insti-

---

172. Kerasidou, *supra* note 167, at 48.

173. Woolley, *supra* note 171, at 112.

174. Angeliki Kerasidou argues that the researcher has to show "good will" similar to that of the treating physician. Kerasidou, *supra* note 167, at 44.

175. Mastroianni, *supra* note 169, at 8.

176. Some even see regulations and sanctions as signs of distrust. *See* Onora O'Neill, *Autonomy and Trust in Bioethics*, 95 J. ROYAL SOC'Y MED. 423 (2002).

tutions and communicating the purpose and value of a treatment or study in a way that patients and study participants can relate to.[177]

V. RISKS OF DATA REGULATION AND PRIVACY LAWS FOR INDIVIDUAL AND PUBLIC HEALTH

Data processing can create risks for privacy and trust, as discussed in the preceding Sections III and IV of this Article. *Not* processing data also creates risks. Restricting the collection, sharing, and other processing of personal data adversely affects the future of medicine and individual health. Such risks resulting from data protection regulations and data privacy laws are often downplayed or overshadowed by concerns for privacy in the public debate regarding privacy laws, but risks resulting from restricting data processing also exist and shall be further illustrated in this section.

A. *Slowing Down Medical and Scientific Progress*

The human body is extraordinary in terms of its complexity.[178] Countless processes take place simultaneously and interact regularly in various ways. Modern medicine has succeeded in understanding many of these processes and has adapted its therapies accordingly. However, without collecting and processing the requisite information, this theoretical knowledge is only of limited use and no further progress is achievable.

Therefore, today's medicine needs more data for treatment, prevention, and medical research. With more data, treatments can become more effective while unnecessary—or even counterproductive—treatments can be avoided.[179] Physicians can get access to an extended pool of known cases via improved data exchange that can be used to evaluate and compare the situation of the patient. Patients also benefit from the use of data processing: for instance, X-ray images, which are stored in an electronic health record and can thus be passed on from doctor to doctor without great effort, are already available when the patient sees another doctor and do not have to be re-created. Treating physicians get a better impression of the current situation and can better tailor their treatment to the individual patient when they have access to a patient's entire medical history.[180]

---

177. Kerasidou, *supra* note 167, at 49.

178. Stephen Naylor & Jake Y Chen, *Unraveling Human Complexity and Disease with Systems Biology and Personalized Medicine*, 7 PERSONALIZED MED. 275 (2010).

179. Nicholas J. Schork points out that every day, a significant amount of people take medication that does not benefit them. *See* Nicholas J. Schork, Comment, *Time for One-Person Trials*, 520 NATURE 609 (2015).

180. Nir Menachemi and Taleah H. Collum name several benefits of electronic health records, e.g., improved legal and regulatory compliance, improved ability to conduct research and increased job satisfaction among physicians. *See* Nir Menachemi & Taleah H. Collum,

Two essential advantages of extensive data processing are comprehensiveness and efficiency. Patient information is essential when treating a disease. Even in the case of specific inquiries, a physician cannot assume that a new patient reliably transmits all the necessary details about personal habits, symptoms, prior treatments, and test results. Often, facts seen as irrelevant are concealed (e.g., drinking or eating habits), exaggerated (e.g., exercise habits), or simply forgotten, even though they are decisive for the correct diagnosis or therapy. For example, knowledge of a low hemoglobin level itself is not necessarily a cause for concern. It could be detected during a routine check-up and would probably not lead to further diagnostic measures, at least not instantly. However, combined with the fact that the patient has been suffering from diarrhea for several weeks, it makes a coloscopy almost indispensable to eliminate more serious diseases as the possible origin of the symptoms. To draw that conclusion, the treating doctor has to be fully informed, which is all but self-evident in a time of increasing specialization. The patient might only focus on specific symptoms based on the special qualification of the respective physician. As a result, a possible connection might be missed.

In addition, time is of extraordinary importance in medicine. As a general rule, the sooner a treatment can be initiated, the better. For example, every tissue in the human body has a certain tolerance for oxygen deprivation. The brain is most sensitive in that regard, a fact that is vividly expressed by the common medical phrase "time is brain."[181] In the case of a stroke every second saved by providing easy access to the complete medical history of the patient could lessen the damage or even save her life.

These advantages have been recognized, but only acted upon in a few countries to date. A good example is the SCAAR registry, which includes all Swedish patients with a coronary intervention.[182] There is no such registry in Germany, for example, which might have something to do with the increased sensitivity and skepticism regarding automated data processing and the more restrictive interpretation and application of data protection laws. Such hesitation, while somewhat understandable in view of the already described dangers of processing health data and that data's particularly sensitive character, stops people from benefitting from the undeniable opportunities.

"Big data" applications can combine and link previously unrelated data sets. This is especially important because of the already-mentioned complexity of the human organism. Traditional research has focused on disease

---

Benefits and Drawbacks of Electronic Health Record Systems, 4 RISK MGMT. & HEALTHCARE POL'Y 47, 47, 50 (2011)

181. Jeffrey L. Saver, *Time Is Brain*, 37 STROKE 263 (2006).

182. Bo Lagerqvist, et al., *Long-Term Outcomes with Drug-Eluting Stents Versus Bare-Metal Stents in Sweden*, 356 N. ENG. J. MED. 1009 (2007).

as a result of one particular physiological change, such as the sudden appearance of particular symptoms or injuries.[183] The physical and mental health of a person, however, is influenced by a variety of factors. Each factor alone can lead to a certain outcome.[184] Once combined with other factors, the effect could be a completely different one. These interactions and correspondences cannot be validated in current clinical trials as they usually focus on the influence of one or two specific variables.[185] Therefore, providing a higher quantity and quality of health data and possibly using computer algorithms to connect and analyze different data sets increases the chances of success of such research and can uncover previously unknown correlations and thus contribute to a better understanding of the causes of diseases and the chances of treatment.[186] Rare but relevant side effects of drugs or therapies can also be detected more quickly and the collection of data on medical treatments makes treatment processes more transparent for the benefit of the patients. Errors can be detected more easily or avoided altogether.[187]

The decoding of the human genome is another factor of enormous importance for the further development of medicine. Genomic data offers a multitude of opportunities for new diagnostics and treatments.[188] Ever since the discovery of the DNA molecule in 1953, the influence of genetics on medicine has been constantly growing.[189] For example, it became a tool in the fight for equality and social justice when the Black Panther Party started to organize testing for sickle cell anemia, an inherited blood disorder that is more common in African Americans,[190] in the 1960s.[191] With regard to personalized medicine in particular, there is an opportunity to develop individual treatment concepts that not only focus on disease as such, but also take into account the patient and all environmental and personal factors affecting

183. *See* Jeffrey. J. Borckardt et al., *Clinical Practice as Natural Laboratory for Psychotherapy Research: A Guide to Case-Based Time-Series Analysis*, 63 AM. PSYCHOLOGIST 77 (2008).

184. Especially on the genetic level, there are numerous variations and seemingly endless possible outcomes. A disease can be the result of an anomaly within one or several genes. Even the specific position of the deviation is important. *See* Soumita Podder & Tapash C. Ghosh, *Exploring the Differences in Evolutionary Rates Between Monogenic and Polygenic Disease Genes in Human*, 27 MOLECULAR BIOLOGY & EVOLUTION 934 (2010).

185. W. Nicholson Price II, *Black-Box Medicine,* 28 HARV. J.L. & TECH. 419, 421 (2015).

186. *See* Price, *supra* note 117, at 1401.

187. For a summary of the benefits of using algorithms in the field of medicine, see *id*.

188. B. M. Knoppers & Yann Joly, *Introduction: The Why and Whither of Genomic Data Sharing*, 137 HUM. GENETICS 569 (2018).

189. *See* Ifeoma Ajunwa, *Genetic Testing Meets Big Data: Tort and Contract Law Issues*, 75 OHIO ST. L.J. 1225, 1227 (2014).

190. *Sickle Cell Trait*, AM. SOC'Y HEMATOLOGY (Feb. 20, 2019, 5:27 PM), http://www.hematology.org/Patients/Anemia/Sickle-Cell-Trait.aspx.

191. Mary T. Bassett, *Beyond Berets: The Black Panthers as Health Activists*, 106 AM. J. PUB. HEALTH 1741, 1741-42 (2016).

her. A good example for optimizing a treatment based on genetic information is the treatment of breast cancer with Trastuzumab, a monoclonal antibody. Certain forms of breast cancer express a particular receptor, the so-called human epidermal growth factor receptor 2 or HER2, that is targeted by Trastuzumab.[192] Therefore, the drug only helps patients whose tumor expresses this receptor, but does that in a very specific and effective way.[193] A particular test can identify the receptor expression rate.[194] In this way, the knowledge about the genetic background of a disease helps researchers and doctors not only to get a better understanding of its pathophysiology but also to find treatments that are more efficient and have fewer side effects, since healthy tissue is less likely to be affected by them. Patients can only fully benefit from these opportunities if laws and regulations do not hinder the collection, use, and exchange of information.

Even from an economic point of view, increased data collection and improved processing can be very useful. It can lead to a higher efficiency and therefore to cost savings in healthcare systems suffering from cost increases worldwide.[195] That can lead to a direct benefit for the contributing patients and the public health insurers.

At the same time, primarily focusing on consent might have unexpected consequences. Health data already is a very valuable good. Further reducing the available amount could result in patients expecting some form of compensation for the requested information, especially in the research context, thereby putting more financial pressure on the healthcare system. And it could cause hospitals or pharmaceutical companies to model their treatment based on whether consent for processing their data has been given by the patients or not.

Data protection law therefore carries the risk of hindering positive developments through the requirements of informed, voluntary, specific, and explicit consent. The general prohibition of automated processing of personal data, the need for data minimization, the obligation to delete data that are no longer acutely needed, and the resulting need for strict purpose limitation in Europe further complicate the process by adding further hurdles to processing a reduced amount of data.

---

192. Gabriel N. Hortobagyi, *Trastuzumab in the Treatment of Breast Cancer*, 353 N. ENG. J. MED. 1734 (2005).

193. *See* Edward H. Romond et al., *Trastuzumab Plus Adjuvant Chemotherapy for Operable HER2-Positive Breast Cancer*, 353 N. ENG. J. MED. 1673, 1674 (2005); *see also* Martine J. Piccart-Gebhart et al., *Trastuzumab After Adjuvant Chemotherapy in HER2-Positive Breast Cancer*, 353 N. ENG. J. MED. 1659, 1660 (2005).

194. *See Breast Cancer HER2 Status*, AM. CANCER SOC'Y (Sept. 20, 2019), https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-her2-status.html#written_by.

195. Price, *supra* note 117, at 140 (referring to precision medicine as an opportunity to "potentially save billions in wasted or inappropriate medical care").

B. *Prevention of Risk*

The collection and processing of data is not only of significant importance for the treatment of diseases, be it on a general level via research or as part of an actual therapeutic measure. It can also be used as a tool for risk prevention. Even now, big data applications are already used this way in the area of public health. The U.S. Center for Disease Control and Prevention ("CDC") is pursuing a strategy of improving data processing and combining data on drug-related deaths to contain and monitor the opioid epidemic.[196] In addition, health authorities around the world use mobile phone data to map population flows in the wake of epidemics or natural disasters, thus anticipating the possible spread of diseases and using their own resources for the greatest possible benefit.[197]

At a more individual level, physicians are required by law to withhold information about an individual's state of health, even if her medical condition could pose a threat to others. In 2015, a pilot flew a commercial aircraft with 150 passengers into a mountain in the French Alps, apparently intentionally, with the intent to commit suicide and mass murder.[198] The pilot had previously received medical treatment from various doctors.[199] On the day of the crash, he had received a doctor's prescription to be on sick leave because of psychological problems.[200] The pilot had kept these circumstances secret from his employer.[201] Due to data protection regulations and patient confidentiality, neither the airline nor the authorities had access to the pilots' health records.[202] According to § 34 of the German Criminal Code, treating doctors are allowed to report a specific health condition and thus to breach the duty of confidentiality in the event of an emergency situation. In this case, however, no reporting took place,[203] possibly due to the uncertainty regarding the correct procedure to follow to ensure compliance with data protection and medical confidentiality laws. Also, in the case of an infectious disease (e.g., an HIV infection), it is not easy for doctors to inform relatives or public authorities if the patient decides to keep his or her disease secret and thereby endangers others.

---

196. *Modernizing Drug Death Data*, CTR. FOR DISEASE CONTROL AND PREVENTION (Apr. 9, 2019), https://www.cdc.gov/surveillance/projects/improving-data-on-drug-overdose-deaths.html.

197. *See* Matthew Wall, *Ebola: Can Big Data Analytics Help Contain its Spread?*, BBC (Oct. 15, 2014), https://www.bbc.com/news/business-29617831.

198. Sven Stockrahm, *War der Absturz vermeidbar?*, ZEIT ONLINE (Mar. 24, 2015), https://www.zeit.de/wissen/2015-03/airbus-a320-germanwings-absturz-frankreich-faq/komplettansicht.

199. *Id.*

200. *Id.*

201. *Id.*

202. *Id.*

203. Strafgesetzbuch [StGB] [German Criminal Code], Nov. 13, 1998, BGBL I at 3322, amended June 19, 2019, BGBL I at 844, § 34 (Ger.).

The protection of a patient's right not to know about health risks in his or her own genetic predisposition has also been a subject of controversy. This right becomes relevant, for example, if the genetic risk of a disease or other health risks are discovered in the course of a clinical study but the discovery is not further processed or shared in a helpful way even though this could have easily been achieved since the information was already collected. This might be another limitation to the collection of relevant information.[204]

If, however, data protection is understood as a self-determined handling of data related to the respective individual itself, it is equally reasonable to assume that the patient has a "right to know."[205] Withholding health information may equal withholding medically indicated treatment or preventive medicine options. The corresponding decision, like the decision on the initial collection of information, is a form of handling personal data, which is the responsibility of the individual. In such constellations, the attending physician is also placed in a position of conflict. The physician has to respect the patient's right not to know but must also act professionally and ethically for the benefit of the patient.

In these cases, a solution based solely on the individual's consent focuses too much on individual privacy and not enough on what is needed to prevent potential dangers to the health of the patient or others. There are good reasons for putting more emphasis on aspects of the public interest and the common good, especially with the increasing inclusion of genetic information in mind. Genetic data is naturally not limited to the individual, but instead touches sensitive data protection concerns of past and future generations along the family line. If, however, the access to or disclosure of genetic information no longer only concerns individual interests, but rather group interests—whether of genetically similar family members or third parties with the same or a similar genetic predisposition and a medical indication—a concentration on individual interests is more difficult to justify.

C. *Law Enforcement and Crime Prevention*

The processing of genetic data can also be highly useful in crime investigations. DNA found at crime scenes can be matched with existing databases to identify or exonerate persons involved.[206] One of the first time this option gained worldwide attention was during the O.J. Simpson trial in the early 1990s.[207] Blood found at the crime scene was identified as being O.J.

---

204. *See* Berkman et al., *supra* note 109, at 57.

205. *See* Schaefer & Savulescu, *supra* note 118, at 22.

206. *See* Zlatko Jakovski et al., *The Power of Forensic DNA Data Bases in Solving Crime Cases*, 6 FORENSIC SCI. INT'L: GENETIC SUPP. SER. 275, 275 (2017).

207. *O.J. Simpson Trial*, ENCYC. BRITANNICA (Jan. 17, 2020), https://www.britannica.com/event/O-J-Simpson-trial.

Simpson's by using then-new DNA testing techniques.[208] His eventual acquittal was seen by some as a result of skepticism of the jury towards this new type of evidence.[209] Today, DNA testing is the gold standard of crime investigation and even after years can help to convict criminals and overturn previous convictions. A recent example is the case of the so-called "Golden State Killer" suspect Joseph James DeAngelo, a former police officer accused of the murder of numerous women in the 1970s and 80s.[210] He was arrested in 2018 after a match between DNA from the crime scenes and a genetic profile belonging to one of his relatives on a website used by people trying to find lost family members was found.[211] Going forward, access to genetic data and automated matching of DNA profiles can make crime investigations easier, faster and more efficient.[212]

VI. POLICY CONSIDERATIONS

The developments in medicine and data protection law described in the preceding sections of this Article call for a reorientation of data protection in the field of healthcare in the twenty-first century. The following policy considerations should be taken into account in the context of developing healthier data protection regulations and privacy laws.

A. *Data Processing Itself Does Not Harm Patients*

Data processing as such does not affect individuals adversely.[213] Individuals can be harmed by inappropriate use of health data, e.g., discrimination by employers or insurance companies. But data processing can also help identify, prove, and counter such inappropriate use through systematic monitoring and analysis, e.g., by applying "big data" capabilities to analyze hiring and contracting practices of employers and insurance companies. Indeed, data processing has many important positive effects on individual health.[214]

---

208. Michael Caruso et al., *O.J. Simpson's DNA Is Linked to the Murder of Nicole Simpson and Ron Goldman in 1994*, DAILY NEWS (Aug. 22, 2016), https://www.nydailynews.com/news/crime/dna-linked-o-simpson-nicole-ron-goldman-murders-article-1.2760781?barcprox=true.

209. *See* Sarah Sloat, *How the O.J. Simpson Trial Created 347 DNA Labs and Public Love for Crime Tech*, INVERSE (Mar. 24, 2016), https://www.inverse.com/article/13129-how-the-o-j-simpson-trial-created-347-dna-labs-and-public-love-for-crime-tech.

210. *See* Matt Stevens, *California Today: How the Golden State Killer Suspect Was Caught*, N.Y. TIMES (Apr. 27, 2018), https://www.nytimes.com/2018/04/27/us/california-today-golden-state-killer-suspect.html.

211. *Id.*

212. *See* Zlatko et al., *supra* note 206, at 276.

213. *See generally* Veil, supra note 132.

214. *See supra* Part V.

Therefore, the general ban on the processing of personal data in EU data protection law should be lifted, either generally or at least in the field of healthcare and medical research and development. If European lawmakers absolutely want to continue with the prohibitive approach of the 1970s, they should provide for broader exemptions for healthcare, medical research, and medical development. The current debate on the role of data protection in medicine tends to focus so heavily on the risks to individual privacy that these risks appear to outweigh any benefits for public and individual health. Yet, this distorted view fails to consider that discrimination and stigmatization are not necessarily consequences of data processing. Instead of regulating data processing, governments should focus laws and enforcement on specific harms and risks.

B. *No One Owns Patient Data*

Neither patients nor businesses should be granted proprietary rights in health data. Calls for data ownership in general,[215] or health information specifically,[216] contemplate that patients should own their health data so they can trade in data property rights by concluding contracts or assigning rights to the data and thus participate in data commercialization and value generation. Property rights in data are not warranted or helpful to promote innovation or other public goods, they would not benefit individuals, and they would suffocate free speech, information freedom, science, commerce and technological progress.[217]

Data protection law aims to control actual access to personal information and - unlike property rights - has neither an incentive nor an investment protection function. Data "belongs" to patients in the sense that they can control access to personal data. The classification of data as economic goods and the resulting commercialization and detachment of information from the individual (ownership can be transferred, the buyer can exclude the

---

215. *See, e.g.*, LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE, 122-35 (1999); Kenneth C. Laudon, *Markets and Privacy*, 39 COMM. ACM 92 (1996); Lawrence Lessig, *The Architecture of Privacy*, 1 VAND. J. ENT. L. & PRAC. 56, 63-65 (1999); Tom C.W. Lin, *Executive Trade Secrets*, 87 NOTRE DAME L. REV. 911, 968 (2012); Patricia Mell, *Seeking Shade in a Land of Perpetual Sunlight: Privacy as Property in the Electronic Wilderness*, 11 BERKELEY TECH. L.J. 1, 26-41 (1996); Richard S. Murphy, *Property Rights in Personal Information: An Economic Defense of Privacy*, 84 GEO. L.J. 2381, 2382-83 (1996); James B. Rule, *Toward Strong Privacy: Values, Markets, Mechanisms, and Institutions*, 54 UNIV. TORONTO L.J. 183, 185 (2004); Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117 HARV. L. REV. 2056 (2004); Catherine M. Valerio Barrad, *Genetic Information and Property Theory*, 87 NW. U. L. REV. 1037, 1062-63 (1993). *But see* Pamela Samuelson, *Privacy as Intellectual Property?*, 52 STAN. L. REV. 1125 (2000).

216. *Digitale Patientenakte bis 2021: Zugriff über eine App / Datenhoheit hat der Versicherte*, FAZ (Sept. 26, 2018), http://edition.faz.net/faz-edition/wirtschaft/2018-09-27/digitale-patientenakte-bis-2021/208239.html.

217. Determann, *supra* note 136, at 1, 42.

seller from continued use) is counterproductive to the objectives of privacy laws, particularly in the context of health information. In addition, linking the use of medical data, which is indispensable for research and development, to financial compensation would not only increase the costs for everyone involved but also further complicate the use of data by research institutions, which would have to be concerned about invisible encumbrances on clinical trial information and individual patient data.

Additional restrictions on access to health data would further hinder the development of new medication and treatment methods as only larger companies and research institutions will be able to "afford" data and might not be interested in sharing it with others. The individual's control of her data is already ensured by existing data protection laws. Administering transactions in data property rights transfers would also necessitate a flood of additional information collection and processing and this would further counter the objectives of privacy laws. No one does or should own data.[218]

C. *Restrictions on Data Sharing Restricts Competition*

Companies are discouraged or even prohibited from sharing personal information with other companies under laws like the GDPR and the CCPA. For example, companies cannot share video footage compiled on public roads to improve child safety for purposes of training autonomous vehicles without providing detailed privacy notices, seeking parental consent, and granting broad access and deletion rights under the GDPR and the CCPA, which make such information gathering impractical for all but the largest organizations. Consequently, innovative start-ups or smaller research institutions lose access to important sources of data that they need to develop competitive treatments, diagnoses, products, and services. Larger organizations, on the other hand, accumulate increased market power and can point to data privacy laws as a reason to deny other organizations access to information as a statutory defense to compliance with competition laws.

D. *Focus on Data Security*

Many risks associated with health data processing can be countered effectively by improving data security. Health data must be stored and transmitted in such a way that both unauthorized access by third parties and the loss of data are prevented. Researchers, physicians, and medical device manufacturers should be encouraged to increase their focus on IT security, make or select more secure products, secure their own data processing systems, and conduct frequent vulnerability tests and audits. HIPAA, CMIA, and a new California law regarding connected devices[219] already contain se-

218. *Id.* at 5, 41.
219. California Consumer Privacy Act, CAL. CIV. CODE § 1798.91.04(a) (West 2020).

curity requirements. Similarly, article 32 of the GDPR and article 22 of the BDSG require organizations to maintain a "level of security appropriate to the risk."[220] This obligation applies broadly, including to individual doctors. Since 2015, larger hospitals are obligated, as "critical infrastructure," under the German IT Security Act to take state-of-the-art measures under state supervision.[221]

Establishing and maintaining a secure data environment requires robust administrative, technical, and organizational measures. Physicians need to implement internal access protocols or staff training, which can cause extra costs and will take time. In the field of information technology in particular, the physician will also often need the help of third parties with the appropriate expertise to fulfill these requirements, even if this outsourcing can lead to increased risks of unauthorized access to health data. From a technical point of view, the realization that complete anonymization of data is practically difficult to achieve should not prevent data from being encrypted.[222] Data security costs money that is currently spent on data minimization and data privacy-related paperwork should be re-applied with a greater focus on data security. More and more sophisticated tools are available.[223]

Effective data security typically requires additional data processing to detect, report, and fight security breaches and perpetrators, to investigate data breaches, and to train artificial intelligence. This is another reason to loosen data collection bans in European data protection law and to qualify or abolish the strict regulations requiring the economic collection of data and their prompt deletion in favor of improved data security.

E. *Introduce General Electronic Health Records and Consents*

In order to use health data more effectively, it is helpful to set up a personal general patient record that is managed independently by a specific institution. Scandinavian countries, which have already established similar systems, could serve as a model in this regard. Swedish patients, for example, can access their complete health records online after setting up an appropriate user account.[224] This could be associated with a right—or possibly a duty—for individuals to decide on the use and provision of their own

220. Council Regulation 2016/679, art. 32, 2016 O.J. (L 119/51-52); Bundesdatenschutzgesetz [BDSG][Federal Data Protection Act], June 30, 2017, BGBL I at 2097, amended by Nov. 20, 2019, BGBL I at 1626, art. 22 (Ger.).

221. Gesetz zur Erhöhung der Sicherheit Informationstechnischer Systeme [IT-Sicherheitsegestz] [Law for Increasing the Security of Information Technology Systems], July 17, 2015, BGBL I, at 1324 (Ger.).

222. *See* Thorogood et al., *supra* note 113, at 1.

223. *See* Price, *supra* note 185, at 1403, 1443.

224. Stephen Armstrong, *Patient Access to Health Records: Striving for the Swedish Ideal*, 357 BMJ 1 (2017).

health data.[225] The individual could thus make an independent decision, which could range from a total refusal to a consent graded by topic and content to a broad consent. Such a one-time and general choice would replace the dozens of consent forms from different institutions.

Another decision on how to deal with the results of processing the data of medically indicated actions, which respects the patient's right not to know, would be necessary. The same applies to the question of how to deal with possible secondary uses of these data.[226]

In certain contexts, particularly in the field of public health, mandatory participation appears to be worth considering.[227] This obligation is based on the idea that patients benefit from improved medical knowledge, which in turn leads to improved health care from which they benefit. In epidemiological research projects on disease control and surveillance in particular, such an obligation could be based on arguments of solidarity, the prevention of "free-riding," and, last but not least, the self-interest of the individual.

General, open-ended consents to keep and use general health records are at odds with current requirements under the GDPR that consent must be specific to be valid.[228] This may be one contributing reason why general health records—despite long-standing plans in Germany,[229] for example—have not yet been realized or made any significant progress in years.

F. *Make Requirements for Voluntary Consent More Flexible*

To prevent patients from feeling like an anonymous object of medicine, a voluntary decision of the person concerned regarding data processing should be promoted as much as possible. Yet it is not necessary that the patient's decision must be documented in a lengthy, incomprehensible form containing all the details required by Articles 12-14 of the GDPR or the CCPA. Patients should be able to declare consent in easy-to-understand, standardized short forms, for example, in the form of a one-time consent concerning access to a general health record. Special provisions for complex situations should be permissible, but specificity should not be legally mandated. The GDPR contains a statutory exception for research purposes,[230] which contemplates that EU member states can exempt data processing for certain medical purposes from GDPR restrictions. It remains to be seen, however, if and how the individual EU member states exercise their discre-

225. Ploug & Holm, *supra* note 142.

226. *See* Thorogood et al., *supra* note 113, at 11.

227. *See* Brent Mittelstadt et al., *Is There a Duty to Participate in Digital Epidemiology?*, 14 LIFE SCI. SOC'Y POL'Y 20 (2018).

228. Council Regulation 2016/679, art. 4(11), 7, 2016 O.J. (L 119/34, 37).

229. *Digitale Patientenakte bis 2021: Zugriff über eine App/Datenhoheit hat der Versicherte*, FRANKFURTER ALLEMEINE (Feb. 20, 2019, 9:55 PM), http://edition.faz.net/faz-edition/wirtschaft/2018-09-27/digitale-patientenakte-bis-2021/208239.html.

230. *See* Mostert et al., *supra* note 94, at 960.

tion and whether organizations will be able to handle the complexities resulting from diverging national legal standards.

For medical research, broad and generic declarations of consent should be regularly accepted.[231] The same problem applies to the use of algorithms or data aggregating apps that compare data to clarify or prevent an unknown medical condition. Overregulating or restricting the use of fitness trackers and similar technology—which is already widely adopted today—is counterproductive. Instead, focusing on developing and improving such technologies should be a policy goal.[232]

New consent models are already being discussed using the terms "broad consent," "open consent," or "dynamic consent."[233] In order to prevent "exhaustion" of the patient through information flooding or to handle the impossibility of authorization by deceased patients, there is a need to standardize and simplify the decision-making process. For this purpose, a general consent at the beginning of using data without explicit feedback and reauthorization would be a good option.[234]

Simplifying the use of health data for research purposes also requires a more flexible approach to the consent of study participants. Long, highly complex consent forms that can hardly cover every conceivable individual case are ineffective and benefit neither the medical practice nor the individual patients. The current legal situation still does not have a concept for granting an "extensive consent." Therefore, changes are necessary, especially in European data protection law.[235] A "sector-specific consent," consent for certain areas of medicine (monitoring of health data in general, treatment of X, research of Y) instead of the common case-specific consent, would simplify data processing for medical purposes. This should be considered for certain particularly trustworthy facilities, especially beneficial uses or relatively harmless data sets.

---

231. *See* Christine Grady et al., *Broad Consent for Research With Biological Samples: Workshop Conclusions*, 15 AM. J.BIOETHICS 34, 35 (2015).

232. William Nicholson Price II sees the use of algorithms as the future of medicine. S*ee* Price, *supra* note 117, at 1401.

233. *See e.g.*, Kristin Solum Steinsbekk et al., *Broad Consent Versus Dynamic Consent in Biobank Research: Is Passive Participation an Ethical Problem?*, 21 EUR. J. HUM. GENETICS 897, 897 (2013).

234. Christine Grady et al. state that participants of the NIH Clinical Center's Department of Bioethics workshop on broad consent agreed that such "broad consent for research use of biospecimens is ethically permissible and, in many cases, optimal." Grady et al., *supra* note 231, at 40.

235. The already described inconsistency between Articles 4(11), 6(1)(a), and 7 of the GDPR on the one hand, *see supra* Part II(2)(b), and Recital 33 on the other hand is especially problematic. Council Regulation 2016/679, recital 33, art. 4(11), 6(1)(a), 7, 2016 O.J. (L 119 /6, 34, 36, 37).

G. *Increase Trustworthiness Through Accountability Certifications*

While patients tend to have trust in data processing such as the creation of a patient record when they see a doctor in person, they tend to be more concerned when data processing is conducted without direct interaction with a trusted physician or other organization. If there is no chance to actually see and supervise the handling of the data, the patient must have confidence in the data processing facility. Trust can be built by implementing comprehensible accountability requirements for organizations to establish a professional and responsible health data processing sector. The evaluation of accountability and compliance requires transparent structures.

Policy and law makers should consider creating or endorsing third party validation systems to which businesses and other organizations can voluntarily submit. Participating organizations could be certified as trustworthy participants and gain easier access to health information. Dual and open certification systems should be considered, including systems certified by governments, e.g., data protection authorities, that could audit the practices of doctors, hospitals, health insurance companies, medical device manufacturers, research institutions, and other organizations. In return for certification, the institutions involved could benefit from a lowering of data protection requirements for the declaration of consent.[236]

Certification would be open to organizations that can demonstrate a particular level of integrity and would involve regular accountability and verification of the documentation of data processing objectives and frameworks with regard to access, use, and control mechanisms. Data processing contrary to the terms of the network would be subject to certification or license revocation and civil and criminal prosecution to ensure compliance with data processing rules and to enhance patient confidence.[237]

In the case of internationally-operating networks, which require data transfer across national borders, there are two options: the network could be designed similarly to the European model of binding rules of conduct, which allow data to be transferred within a group of companies after the underlying regulations have been certified by European data protection authorities. Alternatively, instead of transferring the individual data, the analysis process could be transferred, as it has already been tested under the title

---

236. Recital 33 of the GDPR already contemplates such approaches: "It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognized ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose." Council Regulation 2016/679, recital 33, 2016 O.J. (L 119/6).

237. *See generally* Mittelstadt & Floridi, *supra* note 115.

DataSHIELD.[238] In this context, the data of the participants in a study are stored on individual network computers and processed there according to consistent analysis procedures. The results are then transferred in aggregated and anonymous form so that it is not possible to identify the individual participants. This procedure should result in central merging and processing of the data.[239]

H. *Restriction of Data Subject Rights in the Case of Pseudonymous Data*

The rights of access and deletion of data subjects under current data protection law can cause considerable problems for medical research and endanger the integrity of the stored data. For this reason, data subjects' rights should be limited in the case of research involving health data that does not include the name of the data subject. Similar to the problem of consent to data processing for medical purposes, in many cases the patients will not have expertise and insight into data processing, which prevents a meaningful exercise of rights. The California Consumer Privacy Act already grants an exemption for clinical trial studies and data processing that is covered by HIPAA or CMIA.[240] However, broader exceptions should be considered for other types of research.

VII. OUTLOOK

The future of medicine is shaped by digital transformation, innovative information technologies such as self-learning algorithms (artificial intelligence) and mass data evaluation (big data), personalization, specification, and automated data processing. Moving forward, a treatment considered "state-of-the-art" must be "patient-centric," specifically tailored to the patient's lifestyle, her genetic profile, and as many variables as can be reasonably assessed. This presents a stark difference to medicine of the past, which was more "disease-centric" and focused on the particular illness, its symptoms, and known cures.[241] Data allows treatments to be specific, rather than standardized. In this way, the ideal of the "right drug for the right patient at the right time" could become reality.[242] The success of this transformation depends to a large degree on a healthy balance between data protection regulation, privacy laws, and data access and availability.

---

238. *See* Susan E. Wallace et al., *Protecting Personal Data in Epidemiological Research: DataSHIELD and UK Law*, 17 PUB. HEALTH GENOMICS 149, 151 (2014).

239. *Id.* at 150.

240. California Consumer Privacy Act, CAL. CIV. CODE §1798.145(c) (West 2020).

241. *See* Sherkow & Contreras, *supra* note 114.

242. Wolfgang Sadee & Zunyan Dai, *Pharmacogenetics/Genomics and Personalized Medicine*, 14 HUM. MOLECULAR GENETICS 207, 207 (2005).

Physicians, health insurance companies, scientists, laboratories, medical equipment manufacturers, fitness measuring equipment and service providers, hospital administrations, health authorities, technology companies, and other stakeholders must provide adequate disclosures to patients and other affected data subjects, respect their informational self-determination, and protect their data from unauthorized access and misuse. Failure to do so would result in the affected persons and governments losing confidence in data collection and processing and opting to reject digital improvement that might offer medical progress.

Yet, it is equally, if not more important for individual and public health, that physicians, health insurance companies, scientists, laboratories, medical equipment manufacturers, fitness measuring equipment makers, service providers, hospital administrations, health authorities, technology companies, and other stakeholders have sufficient access to health information. Requirements of EU data protection laws, particularly the EU General Data Protection Regulation, and U.S. privacy laws restricting data sharing, such as the CCPA, threaten to hinder the future of medicine. Undifferentiated data collection bans, a right to be forgotten, excessive requirements for declarations of consent as well as requirements for reduced data use and data deletion slow down development and medical progress.

Excessive data protection can be harmful to health. The German Federal Minister of Health Jens Spahn went as far as saying "[d]ata protection is for the healthy," acknowledging that data protection can become a significant obstacle in the process of treating a disease.[243] Professor Roland Eils, founding director of the Center for Digital Health at the Berlin Institute of Health, warned that excessive data protection threatens lives here and now.[244]

The fear of stigmatization and discrimination in private, at the workplace, or by health insurance companies, has resulted in calls for even stricter data regulations. Yet policy and lawmakers must consider that further restrictions on data processing will also considerably restrict medical progress and obstruct opportunities for the medicine of the future.

EU data regulation focuses on banning or minimizing the collection and use of personal data based on the simple view that data that is not collected in the first place cannot be misused. This approach may have been innovative and worth a try in 1970, when it was first pursued by the German state of Hessen with the declared purpose of preventing George Orwell's vision of 1984 from becoming reality. However, the EU still follows this approach

243. Jens Spahn et al., *App vom Arzt: Bessere Gesundheit durch digitale Medizin*, HERDER FREIBURG (2016).

244. *Mehr als 8.000 Experten beim Hauptstadtkongress*, MEDTECH ZWO (May 23, 2019), https://medtech-zwo.de/aktuelles/nachrichten/nachrichten/hauptstadtkongress-laeuft.html ("Ich würde behaupten, dass ein überzogener Datenschutz jetzt und hier in Deutschland Leben gefährdet.").

today, and is even doubling down on it with the GDPR at a time when this approach seems entirely out of touch with reality. The genie of extensive data collection is long out of the bottle and cannot be put inside again. Data, including sensitive health information, is everywhere. People record it with fitness trackers and smartphones, share it on social media, and leave it on servers of online media companies.

Potential privacy risks do not justify a reflexive call for stricter prohibitions of data collection. Such a reaction does not recognize the opportunities of data processing and improved data exchange for medicine. Instead, it risks turning patients and physicians into adversaries with regard to the collection and processing of health data, despite being largely aligned in regard to their interests. Just as a ban on the production of cars is not appropriate to prevent an increase in road deaths, bans on data processing are not appropriate to fight discrimination or fraud. On the contrary, more and better data collection and processing is needed in the interest of data security, law enforcement, and medical progress.

The ongoing technical development in the field of medicine and the growing focus on the individual as part of more personalized medicine has challenged some traditional data protection principles. The increased possibility of re-identification of anonymized health data and the resulting concerns about dwindling privacy may lead to the anonymization of data being seen as an additional security tool instead of a method of "detachment" from data privacy obligations. However, ways and methods must be found to safely preserve the connection between the data and the data subject for the benefit of the general public and the individual, not dissolve it. With regard to the consent of the individual in the context of medical research, a common good-oriented view highlighting the advantages of participation for the individual should be the preferred way of thinking. With respect to data protection laws, a cautious approach is appropriate. Where, for example, the exercise of data access rights by a person without a background in the medical field results in access to data that the person cannot understand, it does not bring any additional value to the data subject and should not impair the scientific work.

The necessary discussion about the "rules of the game" for the data-based medicine of the future has only just begun. One possible solution is the use of instruments that are already in place in current law. However, an expedient discussion requires the active participation of physicians, who should not leave this future-oriented field to lawyers and politicians alone. Treating and researching physicians are the ones who can explain complex interactions in medical research and build trust. With the future of medicine in mind, in which the protection of personal data is more than an end in itself, their participation in the legal discussion is essential.

The future of medicine offers enormous opportunities and requires a healthy level of data protection. As in drug therapy, the dose makes the poi-

son.[245] When regulation and reduction of the collection and processing of data goes too far, data protection might end up killing more patients than hospital germs.

245. Pracelsus said, "Sola dosis facit venenum" (all things are poison and nothing is without poison; only the dose makes a thing not a poison). F J T Staal, et. al., *Sola Dosis Facit Venenum. Leukemia in Gene Therapy Trials: A Question of Vectors, Inserts and Dosage?*, 22 LEUKEMIA 1849, 1849 (2008).