Open data for DiPH research versus data protection

Merle Freye, Benedikt Buchner

Angaben zur Veröffentlichung / Publication details:

Freye, Merle, and Benedikt Buchner. 2025. "Open data for DiPH research versus data protection." In Digital public health: interdisciplinary perspectives, edited by Hajo Zeeb, Laura Maaß, Tanja Schultz, Ulrike Haug, Iris Pigeot, and Benjamin Schüz, 79–93. Berlin: Springer. https://doi.org/10.1007/978-3-031-90154-6_5.

Open Data for DiPH Research Versus Data Protection

Merle Freye and Benedikt Buchner

Abstract Research in public health is particularly dependent on the unrestricted use of data. The processing of personal health data can serve health-related public interests and bring significant benefits to individuals and society (https://www.oecd.org/health/ministerial-statement-2017.pdf). In contrast, data protection law aims to protect personal data and is consequently based on the prohibition principle and the principle of purpose limitation. Although the General Data Protection Regulation (GDPR) also contains research-friendly approaches and provides the possibility of privileging research data processing at various points, in practice, data protection still poses challenges to research for public health purposes. The concept of consent is questionable, and—especially in Germany—many uncertainties arise from the complex patchwork of state-specific regulations. New concepts such as "broad consent," "dynamic consent," "no consent," and "synthetic data" are intended to facilitate the processing of data for research purposes in line with data protection law. Fueled by these debates, national and international data protection laws will inevitably have to face changes in the near future.

Keywords Anonymization of data · Broad consent · Data protection law · Dynamic consent · GDPR · Medical confidentiality · Synthetic data

M. Freye (✉)
Faculty of Law, Institute for Information, Health and Medical Law (IGMR), University of Bremen, Bremen, Germany

Leibniz Science Campus Digital Public Health (LSC DiPH), Bremen, Germany
e-mail: mfreye@uni-bremen.de

B. Buchner
Faculty of Law, Institute for Bio, Health and Medical Law (IBGM), University of Augsburg, Augsburg, Germany
e-mail: benedikt.buchner@jura.uni-ausgburg.de

H. Zeeb et al. (eds.), *Digital Public Health*, Springer Series on Epidemiology and Public Health, https://doi.org/10.1007/978-3-031-90154-6_5

Abbreviations

BCS70	British Cohort Study
BDSG	German Federal Data Protection Act
DGA	Data Governance Act
DSK	Data Protection Conference
EDPB	European Data Protection Board
EU	European Union
FOG	Forschungsorganisationsgesetz
GDPR	General Data Protection Regulation
GNC	German National Cohort
MII	Medical Informatics Initiative
SOEP	Socio-Economic Panel
StGB	German Criminal Code
UAC	Use-and-access-committees

1 Introduction

As an empirical science, public health is particularly dependent on the unrestricted use of data. Whether it is research into the conditions for health and disease, the interaction between people and their environment, or the evaluation of health system performance: all these public health objectives can only be achieved if as much data as possible are freely available and can be easily linked and tracked over the long term.

The Corona pandemic has emphatically shown us the importance of and dependence on a sufficient database for health sciences. The fact that even after more than 2 years of the pandemic, the Corona Expert Council set up by the German government was still unable to assess the suitability of individual protective measures *due to a lack of data* reveals that Germany is dealing with a structural data problem. During the Corona pandemic, data protection was identified as one of the main culprits for this structural data problem. It is undoubtedly true that the philosophy of Open Data contradicts the basic principles of data protection: in "Big Data," data can "be freely accessed, used, modified, and shared by anyone for any purpose" (FIT4RRI and Foster 2023), on the contrary, data protection law is based on the prohibition principle and strict purpose limitation of data processing.

Nevertheless, the European data protection law is, in theory, not as inflexible as frequently thought since the General Data Protection Regulation (GDPR) was designed to be research-friendly. Accordingly, the GDPR provides special regulations appropriate to research in all central legal questions, which are intended to harmonize the protection of informational self-determination with the special interests and needs of research.

However, the legislator's good intentions cannot overcome the fact that there are still considerable uncertainties in research practice about the application and interpretation of the regulations (Cepic 2021). Data protection and scientific research are

often conflicting (Roßnagel 2019a): On the one hand, "access to, and the processing of personal health data can serve health-related public interests and bring significant benefits to individuals and society […]" (OECD 2017). On the other hand, data protection law aims to protect personal data and thus does not release institutions from their responsibility. Additionally, legal restrictions impeding everyday research may be imposed by national regulations in the member states and consequently challenge the research-friendly setting of the GDPR in practice. Some of these national regulations are general, and others are sector-specific, resulting in an impenetrable jungle of rules. Whether national or international, sector-specific or general—discussion for research primarily revolves around the question of consent. Particularly in large-scale epidemiological studies or genetic studies, the classic concept of (informed) consent is not the most appropriate legal basis (Directorate-General for Health and Food Safety 2021; Chico 2018; Dove 2018). It raises the question of new models of consent and more suitable approaches such as "data donation," "dynamic consent," and "no consent." Although solutions constantly arise, they cannot keep up with technological developments, leading to further uncertainties. For example, the relatively new technical procedure of synthesizing data to generate realistically simulated data from original patient data has recently imposed new challenges for data protection law. The following chapter will discuss whether the current data protection law is sufficient on a national and international level or whether new regulations are inevitable for public health research.

2 Research-Friendly Design of the GDPR

Although data protection was identified as one of the main culprits for the absence of a reliable data basis during the Corona pandemic, this argumentation does not correspond to the concept of GDPR, which is consistently designed to be research-friendly and provides the possibility of privileging research data processing at various points (Buchner 2022).

To begin with, there are exceptions for research projects affecting the two main principles of the GDPR: the purpose limitation principle and the principle of storage limitation. According to the principle of purpose limitation, personal data shall be collected for specified, explicit, and legitimate purposes and not further processed in a manner that is incompatible with those purposes. In deviation from the purpose limitation principle, further processing of data for scientific research purposes is not considered to be incompatible with the original data processing purposes (Art. 5 (1) (b) GDPR). Additionally, the storage limitation principle states that personal data shall be kept in a form that permits the identification of data subjects for no longer than is necessary for the purposes for which the personal data were processed. The exception for research projects is that personal data may be "stored longer" if processed exclusively for research purposes (Art. 5 (1) (e) GDPR).

Moreover, the European legislator enshrined another privilege for research data processing with regard to divergent laws in European countries. According to Art. 9 (2) (j) GDPR, national legislators of the European member states are allowed to

adopt national rules that legitimize the processing of sensitive data in as far as it is necessary for research purposes.

In addition, the GDPR contains exceptions to the rights of data subjects since some rights are not feasible in the research context and would unduly complicate research. When informing the data subject about data processing proves impossible or would involve a disproportionate effort, information shall not be provided (Art. 14 (5) (b) GDPR). For the same reason, the right to erasure does not apply in so far as the right is likely to render impossible or seriously impair the achievement of the objectives of that processing (Art. 17 (3) (d) GDPR). With regard to national law in Europe, member states may provide derogations from data subject rights in their national laws in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfillment of those purposes, Art. 89 (2) GDPR.

Lastly, an essential privilege of research is located in Recital 33 GDPR and introduces the concept of "broad consent." In practice, broad consent is the central issue, dominating the discussion about data protection in research practice.

The given examples emphasize that the GDPR provides special regulations appropriate for research concerning all the central legal issues (consent, legal permissions, general data processing principles, data subject rights), which are intended to balance the protection of informational self-determination with the special interests and needs of research (Buchner 2022).

3 Central Role of Consent

Mainly driven by the exception of broad consent noted above, research for public health purposes in practice is primarily based on consent. In addition to broad consent, another model was developed for research with biobanks: dynamic consent. Consequently, consent plays a central role in research practice. The essential position of consent is not an end in itself since it is also deeply rooted in the ethical principle of autonomy (Beauchamp and Childress 2019): Despite the potential benefits of research for public health purposes to the *public*, the autonomy of the *individual* participant must also be respected. Nevertheless, broad consent and dynamic consent models are not the panaceas they seem to be.

3.1 New Consent Models in Practice…

Consent is one of the lawful grounds on which personal data processing may be based, pursuant to Article 6 of the GDPR. In order to obtain valid consent, certain requirements must be met. Article 4 (11) GDPR stipulates that:

> 'consent' of the data subject means any 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.

Research for public health purposes poses the main challenge that objectives and duration are often not clearly definable or foreseeable at the beginning of the project. Hence, specific information about the planned data processing is not always possible. Broad consent addresses this problem by creating exceptions to the element of "specific" consent. Instead of consenting to specific purposes of data processing, participants consent to broadly defined areas of scientific research or parts of research projects (Raum 2018, marginal no. 35; Caspar 2019, marginal no. 37; Eichler 2023b, marginal no. 2a). Consequently, the exact objective of research projects does not have to be precisely defined from the outset when consent is given; rather, it can be "broader" in the sense of being vaguer.

In Germany, the Medical Informatics Initiative (MII) developed a sample text for a uniform patient information and consent form in close consultation with the Data Protection Conference (DSK) (MII 2020a). The form aims to ensure patient data from care and research are available in an integrated manner at all university hospitals in Germany for direct care and medical research (MII 2020b). Consequently, the sample text affects research for public health purposes that takes part at university hospitals in Germany. Nevertheless, the form has not only gained applause but was also met with severe criticism from the perspective of data protection law (Fröhlich and Spiecker gen. Döhmann 2022). Although this criticism is partly justified, the current version of MII's form is only one national approach to provide guidance on broad consent and is, therefore, not representative of broad consent approaches in Europe. Interestingly, broad consent is also a subject of new European legislative initiatives. Recently, the dazzling notion of "data altruism" in the Data Governance Act (DGA) of 2022 refers to broad consent: According to Art. 2 (16) DGA, so-called data altruism is also based on consent. Recital 50 sentence 5 suggests that data subjects could also consent to certain areas of research or parts of research project. In Austria, broad consent is explicitly named in § 2d (3) FOG (Forschungsorganisationsgesetz), according to which "the indication of a purpose may be made by indicating a research area or several research areas or research projects or parts of research projects ("broad consent")."

In addition to broad consent, dynamic consent has been developed for biobanks. Dynamic consent consists of accompanying permanent information: When new research purposes are determined, new consent forms will be offered (Raum 2018, marginal no. 36; Caspar 2019, marginal no. 41; Eichler 2023b, marginal no. 2d). As soon as the research purpose can be concretized factually and institutionally over time, consent progresses dynamically in parallel (Spitz and Cornelius 2022). Unlike broad consent, dynamic consent thus represents a stretching of the consent process but is not necessarily an exception to specific consent: Essential to this concept is an individual interface or interactive platform tailored to the user's needs. Users can choose how and whether they want to be contacted for new consent and what information they are interested in (Kaye et al. 2015). Consequently, dynamic consent can also integrate broad consent if subjects initially specify that they do not want to be informed dynamically but only once by using broad consent. Dynamic consent has drawn the attention of the German Data Ethics Commission (Datenethikkommission 2019) and the European Data Protection Board (EDPB). According to EDPB, innovative forms of consent in research activities, like dynamic consent, are promising

practices that should be further encouraged and developed (European Data Protection Supervisor 2020).

3.2 …Old Questions

As innovative as broad consent and dynamic consent might appear, both models do not eliminate all the concerns that necessarily accompany consent in public health research projects.

Consent—whether broad or dynamic—is inevitably accompanied by the disadvantage of excluding data processing for which no consent has been given. It is, therefore, questionable whether and to what extent consent-based research data processing can guarantee a database that is representative and free of systematic bias (MII 2020b).

Furthermore, consent comes with a series of deficits that generally call into question the legitimacy of consent as a legal basis for data processing. Regarding the form of MII, subjects still obtain 7 pages of information and, in addition, a 3-page consent text (Sachverständigenrat 2021). Whether and to what extent research subjects are actually (broadly) informed seems at least questionable, given the already very complex daily treatment routine from the patient's point of view. Even dynamic consent only partially solves the information problem. On the one hand, participants obtain information piece by piece, and thus the amount of information per stage decreases. On the other hand, there is no common understanding of informing patients in dynamic consent models, and information thus might be as complex as with broad consent.

When the individuals concerned are patients, the voluntary nature of consent could also be particularly problematic in medical research projects since patients might assume that consent is a prerequisite for treatment. Therefore, the MII Consent Handout states that consent documents should not be presented as part of the admission contract or otherwise suggest to the patient that signing the patient consent would be a prerequisite for treatment. Nevertheless, the MII Consent Handout also recommends asking for consent as early as possible, i.e., before treatment begins or at the start of the treatment process. Hence, the MII Consent Handout itself is not clear on whether, in practice, the consent process is actually "decoupled" from the administrative admission process (MII 2020b). The current concept of broad consent cannot guarantee that patients are not under any perceived pressure to consent. Equally, there is no concept for ensuring the voluntary nature of consent in dynamic consent either.

Another unsolved question relates to the withdrawal of consent. According to the concept of consent, data subjects shall have the right to withdraw their consent at any time, Art. 7 (3) GDPR. Therefore, separate mechanisms must always be provided if data subjects may revoke their consent in the future and thus withdraw the legitimacy basis for further data processing. In practice, this leads to considerable implementation difficulties and jeopardizes a stable data basis for research which is crucial, especially for epidemiological analysis.

Another disadvantage of consent evolves when examining the interaction of consent with the release from confidentiality. Independent from data protection law, Section 203 of the German Criminal Code (StGB) regulates the duty of confidentiality for healthcare professionals. Therefore, a release from the duty of confidentiality must be obtained in order to process health data when a named healthcare professional is involved. This release is different from consent in data protection law. Regarding the data processing by specialized personnel with an obligation of professional secrecy, data protection and medical law lead to a so-called "two-barrier principle" (March et al. 2019; Weichert 2020, marginal no. 138). Consequently, when data processing gains valid consent under data protection law but simultaneously violates a professional secret, the data processing becomes unlawful as a whole (Weichert 2020, marginal no. 146). Although consent under data protection law and medical confidentiality are legally independent, they can still be declared together. This situation potentially overloads patients with information since they must agree twice. Patients may be under the impression that they are asked to make the same legally relevant declaration twice since the requirements of consent and release from the duty of confidentiality are partially similar: they both require the patient's free will and decision, either expressly or impliedly (Bundesärztekammer and Kassenärztliche Bundesvereinigung 2021). While the data subjects shall have the right to withdraw their consent at any time in data protection law, the release from a professional duty of confidentiality can also be revoked at any time with effect for the future (Oberlandesgericht München 2013). Moreover, blanket consents are invalid in both cases (Bundesärztekammer and Kassenärztliche Bundesvereinigung 2021). In conjunction with consent under data protection law, the release from the duty of confidentiality creates a consent bureaucracy that virtually crushes the respective subjects with extensive information. In this situation, it is highly questionable whether patients are adequately informed. Consent plays a central role in research data protection; nevertheless, this primacy should not obscure the fact that consent is not without criticism.

4 Anonymization: Threat or Tool?

Regarding the insufficient concept of consent in the field of research for public health purposes, the question arises whether the solution to the problem lies in anonymizing data. As depicted in Fig. 1, anonymized data must be distinguished from personal data and pseudonymized data.

4.1 Anonymized Data

The GDPR is only applicable to the processing of personal data. According to Art. 4 (1) GDPR, "personal data are any information relating to an identified or identifiable natural person. An identifiable natural person is one who can be identified,

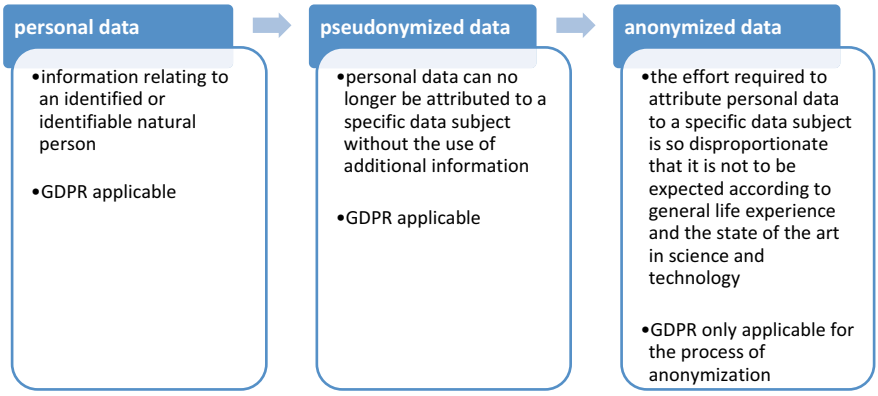

Fig. 1 From personal data to anonymized data. (Source: own figure)

directly or indirectly, in particular by reference to an identifier (such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person)."

Due to data anonymization, subjects become unidentifiable, and consequently, a legal basis for processing this data is not necessary either. Therefore, Recital 26 states that the GDPR does not concern the processing of anonymous information.

Two different theories are discussed in data protection law to decide whether data "relate to an identifiable person." The absolute theory asks if it is impossible *for anyone* to assign an individual information item to a specific person (Schmidt 2022, marginal no. 7). In contrast, the relative theory only considers the means of identification available *to the respective controller in the specific individual case* to establish the personal reference (Roßnagel 2019b, marginal no. 59). The relative theory is not only widely used in academic literature but can also be found in the European Court of Justice decision Breyer v. Germany (European Court of Justice 2016). However, given the constantly evolving technical possibilities, there is no absolute and incontrovertible anonymity of data (Roßnagel 2021). Hence, the idea of absolute anonymization is outdated and does not reflect reality (Hackenberg 2022, marginal no. 53). In contrast to this, the relative approach requires a risk prognosis and determines whether re-identification should be expected according to general life experience and the state of the art in science and technology. This includes the controller's existing or acquirable additional knowledge, current and future technical possibilities for processing, and the possible effort and available time (Roßnagel and Geminn 2021).

With regard to the re-identification risk of research data, it should be noted that the rarer a clinical picture and the smaller a group of subjects, the less anonymity of datasets can be assumed. The re-identification risk is likely to increase significantly, especially in small groups of persons with rare diseases (Conrad and Treeger 2019, § 34, marginal no. 106). Irrespective of this, anonymization does not exempt from

all obligations under data protection law: even if the anonymity of the processed data is assumed, data protection obligations still exist since data protection also protects against possible de-anonymization of data files. Therefore, necessary measures must be taken to reduce this risk as far as possible, including the selection of a sufficiently secure technical procedure for data anonymization and the continuous monitoring and consideration of technical developments (Bundesbeauftragte für den Datenschutz und die Informationsfreiheit 2020).

Although anonymization can be less bureaucratic than obtaining consent in individual cases, in the end, anonymization is still not as straightforward as it seems at first glance. Additionally, the process of anonymization itself constitutes data processing and thus requires a legal basis.

4.2 Pseudonymized Data

Pseudonymization is closely related to anonymization. When anonymization is not possible, the GDPR requires pseudonymization as a legal safeguard for the processing of data for research purposes (Art. 89 (2) GDPR). According to Art. 4 (5) GDPR, "pseudonymization means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

Consequently, the additional information must be kept separately and be subject to technical and organizational measures to ensure that personal data are not attributed to an identified or identifiable natural person. In scientific studies, pseudonymization is convincing when it is essential to be able to contact the individual subject for queries or for information on any medical conditions and diagnoses that may have been determined (Weichert 2022). For example, frequently contacting the participants becomes essential in long-term studies that are based on a follow-up concept, such as the German National Cohort (GNC) or the Socio-Economic Panel (SOEP), and the British Cohort Study (BCS70).

4.3 Synthetic Data

The relatively new technical procedure of synthesizing data has recently raised debates from the perspective of data protection law spinning around the question of anonymization. In 2019, the German Data Ethics Commission recognized the potential of synthetic data—particularly for simulations—and recommended further exploring the generation and use of synthetic data (Datenethikkommission 2019).

In synthesizing data, an AI-supported learning algorithm generates realistically simulated data from original patient data. For this purpose, a dataset's statistical

information and structures are machine-learned in the technical process and "synthesized" with the help of the algorithm (Wegner 2021, Part 6.5, marginal no. 1 ff.). Consequently, personal data are not merely cleansed of identification features according to an anonymization technology, but an entirely new dataset is technically generated based on the trained understanding of an original dataset. With synthesizing, data become more easily accessible to researchers.

From the data protection law perspective, the question arises whether generating synthetic data is an anonymization technique. In this case, the data processing would no longer be subject to the strict limitations of the GDPR due to the lack of personal data processing (Eichler 2023a, marginal no. 27b; Ernst 2021, marginal no. 48). According to the relative approach, a risk prognosis is necessary that focuses on the probability of re-identification risks (Eichler 2023a, marginal no. 15a; Bischoff 2020). To ascertain whether it is likely that the natural person may be identified, all objective factors should be taken into account, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments (Recital 26 GDPR). At present, a risk analysis for synthetic data production is likely to conclude that establishing a personal reference requires too much effort. Following the relative anonymization approach, there would be no personal reference. In order to exclude a personal reference of the AI models, differential privacy could also be considered as an additional step in the generating of synthetic data (Kaulartz and Braegelmann 2020, Chapter 8.9, marginal no. 25). Nevertheless, synthetic data are as yet largely unexplored by legal scholars, and there is no judicial or regulatory practice of application and interpretation. Therefore, no general statement can be made about the anonymity of synthetic data. In individual cases, synthetic data may be personally identifiable, especially in the case of training data that are few in number or highly personalized (Kaulartz and Braegelmann 2020, Chapter 8.9, marginal no. 24).

Synthetic data illustrate that the question of anonymization remains highly contemporary in the end: Whereas new technologies might be a tool for anonymization when they arise, over time, these former novelties rapidly become state of the art, and there might be no absolute and incontrovertible anonymity of data anymore. Consequently, new technologies turn from tools to threats.

5 No Consent

In light of the fact that data anonymization is not always possible and consent is central but not without criticism, the question arises whether data protection knows a path between consent and anonymization that respects the value of public health research for society. According to the GDPR, all processing of personal data is prohibited and requires a legal basis. Besides consent, the European legislature allows other legal bases of equal importance. For example, when processing health data for research purposes, Art. 9 (2) (j) GDPR enables the European member states to

create a legal basis in their national laws and thus offers a legal alternative to consent. In Germany, the legislature made use of this so-called opening clause by creating § 27 of the German Federal Data Protection Act (BDSG): a legal basis for data processing in research without consent. Accordingly, the processing of special categories of personal data shall also be permitted without consent for scientific or historical research purposes or statistical purposes if such processing is necessary for these purposes and the interests of the controller in processing substantially outweigh those of the data subject in not processing the data. This "no consent approach" takes on new relevance as data subjects do not appear to be in favor of consent. According to a Forsa-survey in Germany, only 27% of respondents wanted to be asked for their consent for every new research project (Semler 2019).

Although § 27 BDSG does not require consent and is therefore not subject to the criticism mentioned above, the application of this provision is not trivial either. Instead, the interplay of § 27 BDSG with the data protection laws of the German federal states is highly complex, making it a challenging task to identify the applicable law. So far, Germany's federal and state legislators have failed to create a uniform, practicable, and legally secure legal basis for research data processing. Instead, German law is still characterized by a patchwork of individual regulations that do not reveal any uniform legal concept. For instance, public bodies of the German states can be affected by special laws of the German states, which take precedence over § 27 BDSG. Additionally, there are special laws for hospitals in the federal states and special laws for hospitals under church sponsorship. Due to the legal fragmentation in Germany, it can be complex to determine whether research institutes come under the scope of § 27 BDSG or another law of a federal state.

If § 27 BDSG is applicable, the key question is whether the controller's interests in processing substantially outweigh those of the data subject. On the one hand, this balancing of interest leaves space for the individual circumstances of the specific research project. On the other hand, it leads to legal uncertainty. The fact that a research project concentrates on public health purposes can be a relevant factor in balancing interests. For example, Recital 54 describes that the processing of personal data may be necessary for reasons of public interest in the areas of public health and may be done without consent. Ultimately, it is up to the researchers to ensure that the interests of the persons affected by the data processing are only given a low weight by ensuring the best possible data protection through safeguards and protective measures. For this purpose, researchers have extensive options of guarantees and safeguards, such as anonymization, pseudonymization, and encryption of data, trust and escrow offices, closed data rooms, secure access solutions, and finally also, use-and-access-committees (UAC), which decide on the release of data and ensure compliance with all regulations (Weichert and Krawczak 2019). These and many other instruments not only ensure the highest possible level of data protection in research data processing but simultaneously enable researchers to manage a legally secure balancing of interests in their favor (Buchner 2022). The hurdles for justifying data processing for research purposes without consent are thus not insurmountable.

Besides this "no consent" approach, the opening clause of Art. 9 (2) (j) GDPR allows for a combination of different types of consent with further requirements of another legal basis when processing data for research purposes. In these cases, the primary legal basis is not consent but consent becomes a safeguard. Different types of consent could be the classic informed consent in the GDPR sense but also broad consent as an exception from specific consent. It is important to note that the European Union (EU) countries do not always refer to the same type of consent when discussing data processing for research. Consequently, understanding of consent can vary significantly across EU member states (Directorate-General for Health and Food Safety 2021). For example, the Austrian law (§ 2d (3) FOG) combines broad consent with other prerequisites. In France, consent is required not in a GDPR sense but rather as a national safeguard for the participation of individuals in research (Directorate-General for Health and Food Safety 2021). In the German § 75 SGB X, consent in a GDPR sense is merged with other requirements. Additionally, the buzzword "data donation" (Deutscher Ethikrat 2018; Strech et al. 2020) has started discussions on a new type of legal action in Germany: the so-called "agreement" of the patient. According to the proposal, the agreement is one requirement among others, but it is not equal to consent in GDPR terms since the agreement does not necessarily fulfill the conditions of consent (Strech et al. 2020; Schildbach 2022). Despite all these new consent models, it is not yet foreseeable how the future of data processing for public health research will turn out. To summarize, the GDPR is not only flexible enough to allow for data processing with and without consent but also leaves room for developing new types of consent where the traditional concept of consent no longer appears appropriate.

6 Conclusion

The current legal situation gives rise to considerable practical problems in implementing public health research projects. Although the GDPR theoretically contains several privileges for research, the solutions are still insufficient in practice. Additionally, German law remains a patchwork of individual regulations that do not reveal any uniform legal concept. Partners of joint research projects often belong to different areas with different data protection law requirements, and a distinction must be made between public bodies and private entities. Consent becomes indispensable as the only manageable basis of legitimacy for data processing: To date, practice in public health research is not based on legal privileges but on consent, leading to the fact that consent is central to data processing and perceived as a "silver bullet." However, consent is not without criticism since its conditions are not always fulfilled when the objectives and duration of a project are not definable or foreseeable at the beginning of the project. Although there are approaches for improving consent mechanisms with broad and dynamic designs, the conditions of consent, such as "informed," "voluntarily," and "revocable," are still challenging. Additionally, the interplay of data protection law with medical confidentiality

complicates the consent process. In the end, consent is a stopgap rather than a silver bullet.

Three aspects need attention to overcome these structural difficulties and to realize the potential benefits of public health research for society. First, anonymization processes relating to synthetic data should be further developed and evaluated since they are a promising possibility for processing huge amounts of data, including their scientific reuse. Second, new approaches focusing on a legal basis instead of consent need more attention from international and national perspectives. Whether "no consent approaches" comply with current data protection law is still questionable and therefore demands a clear point of view. Last but not least, the harmonization of research data protection is overdue. Politicians and legislators—at both national and European levels—have put the compatibility of data use and data protection on their agendas and are striving to create a legal framework, particularly for the healthcare sector, that will significantly facilitate data use in the future. In line with the European strategy for data, the EU Digital Governance Act and the EU Data Act aim to make the EU a leader in our data-driven society. Explicitly focusing on the regulation of health data, the European Health Data Space has recently come into force, which aims, among other objectives, to support the use of health data for better healthcare delivery, research, innovation, and policy-making.

References

Beauchamp TL, Childress JF (2019) Principles of biomedical ethics. Oxford University Press, New York

Bischoff C (2020) Pseudoymisierung und Anonymisierung von personenbezogenen Forschungsdaten im Rahmen klinischer Prüfung von Arzneimitteln (Teil I) – Gesetzliche Anforderungen. Pharma Recht 4:309–315

Buchner B (2022) Forschungsdaten effektiver nutzen. Datenschutz und Datensicherheit 9:555–560

Bundesärztekammer, Kassenärztliche Bundesvereinigung (2021) Bekanntmachung: Hinweise und Empfehlungen zur ärztlichen Schweigepflicht, Datenschutz und Datenverarbeitung in der Arztpraxis. https://www.kbv.de/media/sp/Empfehlungen_aerztliche_Schweigepflicht_Datenschutz.pdf. Accessed 10 Feb 2023

Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (2020) Positionspapier zur Anonymisierung unter der DSGVO unter besonderer Berücksichtigung der TK-Branche. https://www.bfdi.bund.de/SharedDocs/Downloads/DE/Konsultationsverfahren/1_Anonymisierung/Positionspapier-Anonymisierung.pdf?__blob=publicationFile&v=4. Accessed 10 Feb 2023

Caspar J (2019) Art. 89 DSGVO. In: Simitis S, Hornung G, Spiecker gen. Döhmann I (eds) Datenschutzrecht: DSGVO mit BDSG. Nomos, Baden-Baden

Cepic M (2021) Broad Consent: Die erweiterte Einwilligung in der Forschung. Zeitschrift für Datenschutz-Aktuell 10:05214

Chico V (2018) The impact of the General Data Protection Regulation on health research. Br Med Bull 128:109–118

Conrad I, Treeger C (2019) § 34 Recht des Datenschutzes. In: Auer-Reinsdorff A, Conrad I (eds) Handbuch IT- und Datenschutzrecht, 3rd edn. C.H. Beck, München

Datenethikkommission (2019) Gutachten der Datenethikkommission. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf;jsessionid=F1C29557CCDB4CB6F3EC21BCC60A379A.1_cid340?__blob=publicationFile&v=7. Accessed 10 Feb 2023

Deutscher Ethikrat (2018) Stellungnahme: Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung. https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-big-data-und-gesundheit.pdf. Accessed 10 Feb 2023

Directorate-General for Health and Food Safety (2021) Assessment of the EU member states' rules on health data in the light of GDPR. https://health.ec.europa.eu/system/files/2021-02/ms_rules_health-data_en_0.pdf. Accessed 10 Feb 2023

Dove E (2018) The EU General Data Protection Regulation: implications for international scientific research in the digital era. J Law Med Ethics 46:1013–1030

Eichler C (2023a) Art. 4 DSGVO. In: Wolff H, Brink S (eds) BeckOK Datenschutzrecht, 44th edn. C.H. Beck, München

Eichler C (2023b) Art. 89 DSGVO. In: Wolff H, Brink S (eds) BeckOK Datenschutzrecht, 44th edn. C.H. Beck, München

Ernst S (2021) Art. 4 DSGVO. In: Paal B, Pauly D (eds) DSGVO BDSG, 3rd edn. C.H. Beck, München

European Court of Justice (2016) Judgment of the Court (Second Chamber) of 19 October 2016 Patrick Breyer v Bundesrepublik Deutschland. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62014CJ0582&from=EN. Accessed 10 Feb 2023

European Data Protection Supervisor (2020) A preliminary opinion on data protection and scientific research. https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf. Accessed 10 Feb 2023

FIT4RRI, Foster (2023) Open and FAIR research data: what is open data and what is FAIR data? https://www.fosteropenscience.eu/learning/open-and-fair-research-data/#/id/5e3741af3ccdf1010dbc6f26. Accessed 10 Feb 2023

Fröhlich W, Spiecker gen. Döhmann I (2022) Die breite Einwilligung (Broad Consent) in die Datenverarbeitung zu medizinischen Forschungszwecken – der aktuelle Irrweg der MII. GesundheitsRecht 6:346–353

Hackenberg W (2022) Teil 15.2 Big Data und Datenschutz. In: Hoeren T, Sieber U, Holznagel B (eds) Handbuch Multimedia Recht, 58th edn. C.H. Beck, München

Kaulartz M, Braegelmann T (2020) Kapitel 8.9 Personenbezug von KI-Modellen. In: Rechtshandbuch Artificial Intelligence und Machine Learning. C.H. Beck, München

Kaye J, Whitley E, Lund D, Morrison M, Teare H, Melham K (2015) Dynamic consent: a patient interface for twenty-first century research networks. Eur J Hum Genet 23:141–146

March S, Andrich S, Drepper J, Horenkamp-Sonntag D, Icks A, Ihle P, Kieschke J (2019) Gute Praxis Datenlinkage (GPD). Gesundheitswesen 81:636–650

Medizininformatik-Initiative (MII) (2020a) Arbeitsgruppe Consent Mustertext Patienteneinwilligung. Version 1.6d as of 16.4.2020. https://www.medizininformatik-initiative.de/sites/default/files/2020-04/MII_AG-Consent_Einheitlicher-Mustertext_v1.6d.pdf. Accessed 10 Feb 2023

Medizininformatik-Initiative (MII) (2020b) Handreichung zur Anwendung der national harmonisierten Patienteninformations- und Einwilligungsdokumente zur Sekundärnutzung von Patientendaten. Version 0.9d as of 16.4.2020. https://www.medizininformatik-initiative.de/sites/default/files/2020-04/MII_AG-Consent_Handreichung_v0.9d.pdf. Accessed 10 Feb 2023

Oberlandesgericht München (2013) Arzthaftung: Widerruf der Einwilligung zur Einsichtnahme in Patientenakte. GesundheitsRecht 8:471–475

Organization for Economic Co-operation and Development (OECD) (2017) Ministerial statement. https://www.oecd.org/health/ministerial-statement-2017.pdf. Accessed 10 Feb 2023

Raum B (2018) Art. 89 DSGVO. In: Ehmann E, Selmayr M (eds) DSGVO, 2nd edn. C.H. Beck, München

Roßnagel A (2019a) Die neuen Datenschutzregelungen in der Forschungspraxis von Hochschulen. Zeitschrift für Datenschutz 4:157–164

Roßnagel A (2019b) Art. 4 DSGVO. In: Simitis S, Hornung G, Spiecker gen. Döhmann I (eds) Datenschutzrecht: DSGVO mit BDSG. Nomos, Baden-Baden

Roßnagel A (2021) Datenlöschung und Anonymisierung. Zeitschrift für Datenschutz 4:188–192

Roßnagel A, Geminn C (2021) Vertrauen in Anonymisierung. Zeitschrift für Datenschutz 9:487–490

Sachverständigenrat (2021) Gutachten: Digitalisierung für Gesundheit – Ziele und Rahmenbedingungen eines dynamisch lernenden Gesundheitssystems. https://www.svr-gesundheit.de/fileadmin/Gutachten/Gutachten_2021/SVR_Gutachten_2021.pdf. Accessed 10 Feb 2023

Schildbach R (2022) Zugang zu Daten der öffentlichen Hand und Datenaltruismus nach dem Entwurf des Daten-Governance-Gesetzes. Zeitschrift für Datenschutz 3:148–153

Schmidt B (2022) Art. 2 DSGVO. In: Taeger J, Gabel D (eds) DSGVO – BDSG – TTDSG, 4th edn. dfv, Frankfurt am Main

Semler S (2019) "Datenspende" für die medizinische Forschung: Ergebnisse einer Forsa-Umfrage im Auftrag der Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V. (TMF). https://www.tmf-ev.de/News/articleType/ArticleView/articleId/4456.aspx. Accessed 10 Feb 2023

Spitz M, Cornelius K (2022) Einwilligung und gesetzliche Forschungsklausel als Rechtsgrundlagen für die Sekundärnutzung klinischer Daten zu Forschungszwecken. Medizinrecht 40:191–198

Strech D, Kielmansegg S, Zenker S, Krawczak M, Semler S (2020) Wissenschaftliches Gutachten „Datenspende"– Bedarf für die Forschung, ethische Bewertung, rechtliche, informationstechnologische und organisatorische Rahmenbedingungen. https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/5_Publikationen/Ministerium/Berichte/Gutachten_Datenspende.pdf. Accessed 10 Feb 2023

Wegner S (2021) Teil 6.5 Synthetische Daten. In: Leupold A, Wiebe A, Glossner S (eds) IT-Recht – Recht, Wirtschaft und Technik der digitalen Transformation, 4th edn. C.H. Beck, München

Weichert T (2020) Art. 9 DSGVO. In: Kühling J, Buchner B (eds) DSGVO BDSG, 3rd edn. C.H. Beck, München

Weichert T (2022) Datenschutzrechtliche Rahmenbedingungen medizinischer Forschung: Vorgaben der EU-Datenschutz-Grundverordnung und national geltender Gesetze. MWV Medizinisch Wissenschaftliche Verlagsgesellschaft mbH & Co. KG, Berlin

Weichert T, Krawczak M (2019) Vorschlag einer modernen Dateninfrastruktur für die medizinische Forschung in Deutschland. GMS Medizinische Informatik, Biometrie und Epidemiologie 15(1):Doc03