Clinical trials data transparency & GDPR compliance: What are the effects on data sharing and open innovation?

Minssen, Timo; Rajam, Neethu; Bogers, Marcel

Published in:
Science and Public Policy

DOI:
10.1093/scipol/scaa014

Publication date:
2020

Science and Public Policy, 47(5), 2020, 616–626
doi: 10.1093/scipol/scaa014
Advance Access Publication Date: 4 March 2020
Article

Clinical trial data transparency and GDPR compliance: Implications for data sharing and open innovation

Timo Minssen [1,*], Neethu Rajam [1] and Marcel Bogers[2,3]

[1]Centre for Advanced Studies in Biomedical Innovation Law (CeBIL), University of Copenhagen (UCPH), Copenhagen, Denmark, [2]Department of Food & Resource Economics (IFRO), University of Copenhagen (UCPH), Copenhagen, Denmark and [3]Haas School of Business, University of California, Berkeley, CA, USA

*Corresponding author. E-mail: timo.minssen@jur.ku.dk

Abstract

Recent European Union (EU) initiatives and legislation have considerably increased public access to clinical trials data (CTD). These developments are generally much welcomed for the enhancement of science, trust, and open innovation. However, they also raise many questions and concerns, not least at the interface between CTD transparency and other areas of evolving EU law on the protection of trade secrets, IPRs, and privacy. This article focuses on privacy issues and on the interrelation between developments in transparency and the EU's new General Data Protection Regulation 2016/679 (GDPR). More specifically, this article examines: (1) the origins and rationales of EU transparency regulations, including the incidents and concerns that have shaped them; (2) the features and implications of the GDPR which are relevant in the context of clinical trials; and (3) the risk for tensions between the GDPR and the policy goals of CTD transparency, as well as implications for data sharing and open innovation. Ultimately, we elaborate on factors that should be carefully considered and addressed to reap the full benefits of CTD transparency.

Key words: Clinical Trial data, GDPR, transparency, open innovation

1. Introduction

Clinical trials are critical to our understanding of the safety and efficacy of medical innovation and the products which it gives rise to (Lexchin et al. 2003). Until recently, however, much of the information resulting from clinical trials remained confidential and only the results from clinical trials were openly accessible. This has changed fundamentally. The new US/European policy initiatives and new legislation, such as the European Medicine Agency (EMA)'s Policy 0070[1] and the European Union (EU)'s new clinical trial regulation (CTR) 536/2014,[2] have considerably increased public access to clinical trials data (CTD). The new disclosure rules not only encompass the results of clinical studies, but also pertain to anonymized patient-level data and other detailed information from clinical trials' dossiers. In simple terms, CTD transparency implies that decisions and data from clinical studies are widely shared with other researchers, clinicians, and the public (Institute of Medicine 2015). These new initiatives are generally perceived as much welcomed developments for the enhancement of science, scientific collaboration, trust, and open innovation.

Yet, while in line with the increased openness that has characterized innovation in the private and public sector (Bogers et al. 2018), these developments also highlight concerns about how to balance the 'intrinsic tensions between closed and open innovation' (Martin 2016). In this respect, potential drawbacks and concerns associated with increased transparency also need to be carefully considered. Conducting clinical trials is a lengthy, risky, and expensive process involving huge investments with often low success-rates. Hence, it comes as no surprise that some commentators have expressed concerns over fewer incentives to invest in innovative research and development as the expansion of public general knowledge would result in barriers to intellectual property and the protection of trade secrets (Rajam and Minssen 2019), and to an increased exposure to litigation following from the systematic examination of such data (Price II and Minssen 2015). Others have expressed additional worries over patient privacy and the possibilities for misuse of CTD (Mello et al. 2013; Institute of Medicine 2015).

This article focuses on selected legal issues relating to privacy and personal data within the context of European CTD transparency and discusses their potential implications for the European open innovation landscape. A discourse of this nature is very timely due to the increasing importance of transparency and openness for the European innovation system. Many of the privacy and personal

data-related worries in the new Big Data era have been addressed in seminal new legislation, the EU General Data Protection Regulation (GDPR) 2016/679.[3] The primary goal of the GDPR, which applies to EU and foreign companies, authorities, and other stakeholders who process the personal data of subjects residing in the Union, is to install an EU-wide mechanism for enhancing and strengthening the protection of personal data. Non-compliance with GDPR stipulations can result in substantial fines and therefore understanding and complying with the GDPR requirements can be a daunting task. This has resulted in a particularly challenging situation for small and medium-sized organizations and companies, but even larger bodies, such as the EMA, have to devote considerable resources to ensuring GDPR compliance and the development of an effective system for the anonymization of patient-level data. But, to what extent does this correspond with the underlying rationales of data transparency? How might complex GDPR requirements influence the EU's high-level goals on open innovation? And if so, in which circumstances could redactions or publications be justified when balancing values and interests?

With these questions in mind, this article aims to describe the genesis of the regulatory transparency requirements as well as the incidents, developments, and policy concerns which have shaped them. Special emphasis will be placed on the interface between the EU's new CTR, the EMA's transparency policy and the GDPR. More specifically, we will examine legislative shortfalls and the risk for tensions between GDPR requirements and the high-level EU policy goals for openness and data sharing.

2. The significance of CTD transparency for research and open innovation

Data in any form are essential for biomedical research, clinical trials included (Travis 2011; Toga et al. 2015; Packer 2018) Clinical trials are lengthy, costly, and rigorously controlled tests designed to examine the safety and effectiveness of medicines, devices, treatments, or preventive measures in humans (Martin et al. 2017).[4] By assessing the pharmacological, pharmacokinetic, and /or the effect of other pharmacodynamics of the investigational agent or any adverse reaction of such agent, clinical studies and CTD provide vital information on the quality, safety, and efficacy of drugs (Mohs and Greig 2017). With more transparency, independent researchers can more effectively enact checks on sponsors by verifying results in order to detect important findings on product safety and effectiveness (Mello et al. 2013). Independent reanalysis of clinical study reports can also inform regulatory authorities of the safety and effectiveness of products that are already on the market.[5] Moreover, aggregated participant-level data enables scientists to examine questions that have not been sufficiently addressed in individual trials and to find new uses of known drugs (Mello et al. 2013). These conclusions and discoveries are especially relevant for research in public health and personalized medicine. The availability of information also enables responsible research in furthering the researcher's commitments to the participants. In addition, transparency can inform new directions for different kinds of research and investment on the part of pharmaceutical companies and thus represents an indispensable tool for clinical research which simultaneously creates new opportunities for open innovation.

Various forms of open innovation, such as open-source, innovation centers or crowdsourcing, have been identified as contributing to innovation potential in research-based pharmaceutical companies (Schuhmacher et al. 2016). Within this trajectory, data sharing and transparency will require novel research approaches (Longo and Drazen 2016) and governance models (Hudson et al. 2016). While this may be linked to a new paradigm in pre-competitive pharmaceutical research (Bhardwaj et al. 2011), such openness would ultimately integrate data transparency and availability of information into a business model that is based on purposive management of knowledge flow across organizational boundaries (Chesbrough and Bogers 2014). Evolving such a model would require dealing with the underlying tensions between sharing and the protection of this knowledge (Bogers 2011) while acknowledging the risks of a more open model (Lowman et al. 2012). By aiming to foster innovation and research, and in avoiding unnecessary duplication of clinical trials or repetition of unsuccessful trials, CTD transparency can contribute to enhancing the efficiency of large collaborative clinical trials, with the greatest benefit for those conducted collaboratively across multiple Member States. Thus, well-administered CTD transparency could give rise to enormous benefits, inter alia for the enhancement of precision medicine, public trust, and open innovation.

3. Recent developments and emerging complications in transparency

Global and European efforts have long been underway from campaigners, researchers, and patient groups to increase data transparency for clinical trials (Kmietowicz 2014). Since 2010, the EMA has developed policies to release documents based on transparency regulations.[6] Its policy on the transparency of clinical data rests on the belief that it would reinforce public trust and confidence in the EMA's scientific and decision-making processes in addition to preventing duplication of clinical trials and promoting innovation in the area of new medicines.[7]

The drug industry has, for multiple reasons, supported some of these developments and has even responded with its own transparency projects, including initiatives by GlaxoSmithKline, AstraZeneca, Sanofi, and Pfizer (among others) (Price II and Minssen 2015). Even though compliance with the new transparency requirements is increasing (Lassman et al. 2017), the disclosure of clinical data by the EMA has also led to opposition from the pharmaceutical industry in granting access to specific information submitted in the dossier of an application for marketing authorization (MA) for a medicinal product.[8] Opponents of disclosure often contend that much of the information contained in clinical trial reports is covered by a general presumption of confidentiality. This opposition stems partly from concern over the transmission of valuable data to competitors and partly from the obligation to maintain confidentiality of different kinds of information. In this regard, it is important to remember that in addition to commercially sensitive business information, CTD encompasses personal data pertaining to different participants such as patients, personnel, sales, sub-contractors, etc. Clinical trial sponsors are thus also controllers of data and have a responsibility to maintain the confidentiality obligations provided for by laws such as the GDPR (Mello et al. 2013). Therefore, companies will not only refer to the protection of their intellectual property rights (IPRs) and trade secrets,[9] but also to the GDPR when arguing for the non-disclosure of data sets.

It will come as no surprise that the EMA's transparency policy has led to several proceedings before the European Court of Justice of the EU (CJEU).[10] So far, these proceedings have primarily concerned the issues of protecting commercially confidential

information and trade secrets,[11] but it is easy to imagine how personal data protection rules will be evoked in similar future proceedings.

The relevance of these decisions has increased since 2014, when the EMA amplified its data-sharing policies through the aforementioned Policy 0070,[12] which removed access restriction to allow researchers to download, save, and print clinical study reports for academic and non-commercial research purposes (Bonini et al. 2014; Watson 2014). Moreover, the adoption of the EU Clinical Trials Regulation (CTR) brought in significant changes in terms of the transparency scenario. The CTR provides that all relevant information regarding the clinical trial be submitted through the publicly accessible EU portal. It is hoped that publicly available information contained in the EU database will contribute to protecting public health and fostering the innovation capacity of European medical research, while recognizing the legitimate economic interests of sponsors.[13] However, while offering access to the data and information contained in the public database, confidentiality is respected where it is essential for protecting:

- personal data
- commercially confidential information
- confidential communication between member states
- or ensuring effective supervision for the conduct of clinical trials.[14]

This provision under the CTR also falls in line with another EU Regulation on public access to documents.[15] The CTR requires that no personal data of subjects shall be publicly accessible.[16] The EU database shall contain personal data only insofar as this is necessary for enabling cooperation between the competent authorities, to provide information about specific clinical trials, previous submissions, or communication of clinical information about medical products to the public, and in other similar circumstances.[17] In ensuring personal data protection, the EMA has also been engaged in several efforts including the drafting of guidelines for the publication of clinical data for medicinal products for human use,[18] and an expert group on data anonymization known as the technical anonymization group. Such efforts are aimed at developing best practices for the anonymization of clinical reports, that is, in the context of the EMA's policy on the publication of CTD.[19] With regard to the anonymization of personal data, the EMA's Guidelines follow three main criteria, namely:

a. Possibility to single out an individual.
b. Possibility to link records relating to an individual.
c. Possibility to make inference concerning an individual.

In addition, where such criteria are not met, the EMA ensures an assessment based on the risks involved. The EMA has also developed different measures for managing the disclosure requirements for data containing personal information which may involve retaining the data, redacting, or removing it from the publicly available database. These measures are supposed to ensure that the disclosure policy of the EMA offers adequate personal data protection.

A report published in July 2018, revealed that during the first year of implementation (October 2016–7) of the clinical data publication policy (Policy 0070), the EMA had published clinical reports for ~50 medicinal products under this framework.[20] The report provides a detailed picture of how the data are disclosed, including information on who had access to the data and how it was processed following from the request for access.[21] However, these early disclosures did not yet include raw data or individual patient data from the clinical trials.

The EMA has, however, planned for the second phase of implementation of its Transparency Policy, during which it aims to publish the individual patient data contained in these clinical trial reports.[22] Providing access to the individual patient data collected in clinical trials furthers research that may advance medical science or improve patient care. In turn, this helps to ensure that the data provided by research participants are used to maximum effect in the creation of new knowledge and understanding (Hughes et al. 2014). Although there are clear benefits to providing greater access to individual patient data, a number of aspects need careful consideration. These include providing access in ways where the risks to patient privacy and confidentiality are minimized, and the commitments made to patients via informed consent processes are adhered to.[23] This calls for greater attention to issues of personal data protection (particularly consent), data transfer, and the use of publicly available data.

4. The interface between GDPR and clinical trials

One of the primary goals of the GDPR is to safeguard the rights of individuals to have reasonable control and be better informed about how their data is being used. To achieve this, the GDPR imposes new roles and obligations on data controllers and processors along with a stronger legal foundation for processing personal data.

The GDPR's definitions of the terms 'processing' data as well as 'personal data' are broad. They cover health data, such as health records and CTD, even if they do not directly identify a person or allude to their health status (Price II et al. 2019).[24] Under the GDPR, data concerning health[25] is considered to be sensitive data, the processing of which is forbidden by default.[26] In order to process such sensitive health data and biometric data, companies, and organizations will either need to gain explicit consent from relevant persons, or assess whether they fall under various GDPR exceptions to this ban. These include, *inter alia*, exceptions for medical treatment, and for the 'public interest in the area of public health'.[27] But in such cases, companies and organizations may need to identify other laws that give them permission and will need to institute additional safeguards when processing occurs (Price II et al. 2019).[28] While the GDPR leaves room for individual EU countries to enact laws governing health information or leave their own more restrictive laws in place, many of these exceptions apply only where the EU or a Member State has enacted a relevant law.[29] Therefore, healthcare professionals, companies, authorities such as the EMA, and other stakeholders 'processing' the 'personal data' of subjects residing in the Union must comply with the standards and duties stipulated in the regulation even if individuals do not invoke their rights.[30]

Clinical trial providers and authorities might not be processors alone; they may also be considered as data controllers with regard to the clinical trial procedures and therefore must ensure that the applicable procedures are in place and rules are followed.[31] Therefore, clinical trial providers must identify what data are being processed, who processes the data, where it is transferred to, what it used for, any risks related to the processes, and finally ensure that all employees are trained to meet these objectives (Gogates 2018). It is also the duty of the clinical trial sponsors who may also be controllers of the clinical data to maintain records of data processing activities and perform data processing impact assessments in the interests of protecting the rights of clinical trial participants.[32] This obliges the

controller of clinical data to ensure that the interests and fundamental rights of the data subject are in no way affected, particularly in circumstances where said data undergo further processing.[33] The GDPR does not only apply to clinical trial participants, but to employees, customers, and subcontractors as well. The GDPR being extensive in its definition of personal data, therefore, has wider implications for clinical trials in terms of participants and other stakeholders involved in the process.

The European Data Protection Board (EDPB) recently adopted an Opinion (3/2019) on the interplay between the EU CTR and the GDPR. The opinion notes that both the GDPR and CTR expressly refer to each other and hence it follows that both legislations apply simultaneously and that the CTR constitutes a sectoral law containing specific provisions that are relevant from a data protection viewpoint without derogating from the GDPR.[34] The opinion distinguishes primary and secondary legal bases for the processing of personal data for clinical trials. The primary legal basis for processing CTD under the GDPR relates to reliability and safety purposes, while the secondary legal basis for processing purely relates to research activities pertaining to CTD. [35] This will require organizations involved with clinical trials to evaluate the nature of data processing to ascertain its legal basis under the GDPR. However, there are specific interrelated challenges associated with this evaluation process which will be elaborated upon in the next section.

5. GDPR and selected challenges to CTD transparency

In the following, we will focus on four selected challenges at the interface of the GDPR and CTD transparency, namley: Consent, data uses, anonymization, and international data transfer.

5.1 Consent

Data Protection and Good Clinical Practice (GCP) are essential components of clinical research. These have been regulated to establish a harmonized approach to clinical research and have evolved with inputs from the OECD Guidelines for Data Protection and the ICH E6 Guidelines for GCPs.[36] Furthermore, 'human dignity' and the 'right to the integrity' of the person are recognized in the Charter of Fundamental Rights of the European Union (the 'Charter'). In particular, the Charter requires that any intervention in the field of biology and medicine not be performed without free and informed consent of the person concerned.[37] Consent is therefore highly important for clinical trials.

Under the CTR, clinical trials can only be conducted where informed consent for the processing of personal health data has been given. Where a subject is not able to give informed consent themself, their legally designated representative must be informed and called upon to give consent on their behalf.[38] Such consent shall be written, dated, and signed by the person giving the consent or may, in special cases, be recorded. Where the clinical trial relates to a vulnerable subject, additional expert opinions may be sought as authorization for the trial.[39] The subject, or their legally designated representative, may withdraw the consent given for clinical research at any time.[40] The GDPR also requires the controller to demonstrate that the data subject has consented to the processing of their personal data.[41] Health data, being 'special categories of data', can only be processed if specific conditions are met.[42] However, the notion of 'consent' under the GDPR is essentially different from the notion in the CTR.[43] Under the CTR, informed consent means a subject's free and voluntary expression of their willingness to participate in a particular clinical trial after having been informed of all aspects of the clinical trial that are relevant to the subject's decision to participate.[44] At the same time, under the GDPR, consent is given for the processing of personal data. Further, under recital 161 the GDPR states that for the purpose of 'consenting to the participation in scientific research activities in clinical trials, the relevant provisions of Regulation (EU) No 536/2014 of the European Parliament and of the Council should apply'.

Yet, consent is not always required to process data under the GDPR, especially if it is for the purpose of scientific research, subject to appropriate safeguards.[45] The EU Guidelines on consent issues[46] observe that data processing carried out in order to gather reliable and robust data in line with a clinical trial protocol as approved by the competent regulatory authorities under the CTR may be considered 'necessary for compliance with a legal obligation', and so lawful under the GDPR without the need to obtain consent from patients (Harford et al. 2018). However, the EDPB opinion calls for different consent for different types of clinical data processing which may involve a multiplicity of consent forms which may add to the complexities of the clinical trial process and the understanding of its participants. Furthermore, where an opportunity is not provided to give consent to different data (Kaye et al. 2015) processing separately, this should not automatically imply unlimited consent for the purposes of processing of data nor should it create an imbalance between the data subject and the controller of data.[47]

In this respect, clinical trials must be carefully monitored and must be subject to ethics and GDPR impact reviews particularly view of the issue of consent and any imbalance that it may create.[48] Good practices of patient involvement in trial prioritization and design, as well as the use of new technologies allowing for more 'dynamic informed consent approaches', could be a starting point to ensure that clinical trials and follow-on procedures meet the standards set for CTD processing.[49]

5.2 The use of publicly available data

The protection of personal data aims to ensure compliance with the fundamental rights of EU citizens. The EU regulation (EC) No 45/2001 along with Directive 95/46/EC, also offers different means to anonymize data and protect patients from retroactive identification. Yet, emerging technologies for data mining and database linkage increase the potential for unlawful retroactive patient identification and the use of such data.[50] It is therefore essential to address the use of publicly available data. In that regard, it is important to stress that under Article 3 (2) GDPR the GDPR also applies to the processing of the personal data of data subjects residing in the Union by a controller or processor not established in the Union, where the processing activities are related to:

a. the offering of goods or services, irrespective of whether a payment to the data subject is required, to such data subjects in the Union; or
b. the monitoring of their behavior provided that their behavior takes place within the Union.

Where the entity processing the document may not be subject to the GDPR, it may still be used in different ways. However, this use may be problematic from a privacy perspective if it relates to processing of 'personal data'.

In the context of clinical trial reports, where the data are processed up to at least two levels with anonymization, redaction, and

removal, the data received by third party users of the CT database would, at best, amount to 'research data'. And where databases offer merely *pseudonymized* data of a data subject within the EU, its access and transmission to third parties are limited by the scope of the GDPR.[51]

However, what if publicly available data is analyzed by companies outside the EU and then transmitted to third parties outside the EU? This onward transfer of data would normally count as processing under the GDPR. But if this processing is not related to activities set out in Article 3(2) GDPR it would go beyond the GDPR's scope. Additionally, foreign activities that might involve the use of new technologies that might be used to 're-identify' anonymized research data are difficult to monitor and the third party company buying such data could also fall outside the scope of the GDPR if such activities would not fall under the specified conditions.[52]

Public data can be of different types depending on their origin, such as (1) data actively and knowingly provided by the data subject, or (2) observed data provided by the data subject by virtue of the use of the service or the device (for instance, through the use of medical devices).[53] In addition, there could be a third category of data that is created by the data controller on the basis of the data 'provided by the data subject'. Such data may be called inferred, derived, or analyzed data.[54] Clinical reports involve clinical overviews, clinical summaries, and clinical study reports reflecting a number of pharmacological or other pharmacodynamic effects on patients within the clinical trials.[55] These require derived data and analysis of data from subjects which, though personal, may be subject to the GDPR. C-141/12[56] holds that a legal analysis of an individual is not 'in itself' personal data even though it contains personal data.[57] Although the GDPR provides a right of access to such data, its exploitation by third parties, particularly in the aforementioned scenarios where the GDPR has geographic limitations, leaves possibilities for unwarranted uses of CTD. According to the EMA's conditions of use, the user of clinical trials reports:

1. May use the report for general information and other non-commercial purposes, including non-commercial research purposes, subject to these terms.
2. May not use the clinical reports to support an application to obtain an MA and any extensions or variations thereof for a product anywhere in the world;
3. Is not granted any intellectual property or other commercial rights in relation to the clinical reports other than as expressly set out in these terms.[58]

However, given that data including derived data are publicly transmitted under the transparency rules, preventing the misuse of clinical data in unanticipated ways poses a challenge for devising appropriate controls against this phenomenon.

5.3 Anonymization

CTD encompasses different types of data such as clinical reports, patient-level data, and other processing data. As a rule, under the transparency policy, all clinical reports submitted as part of a regulatory application will be subject to publication and may even be subject to third party access.[59] In this process, the EMA's transparency policy establishes methods for balancing the protection of patients' privacy through the anonymization of protected personal data. This involves a process of rendering data into a form which does not identify individuals and where identification is not likely to take place.[60] Certain types of clinical information may also have to be de-identified which involves a process of removing or obscuring any personally identifiable information from individual records so as to minimize the risk of unintended disclosure of the identity of individuals and information about them. Different strategies exist for anonymization and de-identification of personal data, such as deletion, redaction, generalization, perturbation, or dissociation of identifying information.[61] Clinical transparency must thus ensure that the data in clinical reports must be processed in a way that ensures a reasonable degree of anonymization by the data controller.[62]

The EMA's Policy 0070 requires the anonymization of clinical reports not be subject to redaction alone as this tends to decrease the clinical utility of the data.[63] Therefore, it is essential that different techniques are applied to the anonymization of data. However, the appropriateness of the different processes of anonymization is left to the discretion of the data controller. This poses a challenge as different anonymization strategies have been dealt with differently by different scientific experts and bodies (see for instance Dias 2015; PhUSE De-Identification Working Group 2015).[64] Such an anonymization exercise must address the problem of singling out, linkability or inference, which offers a possibility to identify personal data.[65] In addition to assessing the appropriate anonymization technique, the data controller must also ensure that there is no risk of re-identification through reasonable means.[66] Further, for publication of clinical reports the EMA's Policy also requires that the sponsor submit an anonymization report along with a justification of the redactions for commercially confidential information.[67]

This could be even more problematic in relation to individual patient data, which the EMA plans to release in the next phase of its transparency efforts. Individual patient data containing quasi-identifiers, which consist mostly of dates, location information, demographics, socioeconomic information, rare diagnoses, concomitant illnesses, and medications, and serious adverse events such as death, hospitalization, rare diseases, and birth defects, cannot be removed as these variables are very useful for analysis (El Emam and Aballah 2015).[68] These variables require more sophisticated techniques for retaining their informational value whilst reducing the probability that these variables can re-identify participants through data triangulation (see also Price II and Cohen 2019; Cohen and Mello 2019). These data assessment processes are very time consuming and do not ensure patient confidentiality completely (Tucker et al. 2016). The increase in publicly available data over time, coupled with the de-identified CTD and new technologies, such as quantum computing (Donovan 2018), may increase the risk of patient re-identification (Tucker et al. 2016). Ensuring anonymization of such CTD may require substantial technical measures which tend to increase the effort required as well as the cost of making patient-level data available for third party research (Tucker et al. 2016).

5.4 International data transfer

The ability to access, use, and share information across borders stimulates innovation, and in particular clinical trials. Therefore, limitations on cross-border data flow in the clinical trials sector presents serious challenges (Van Der Marel et al. 2014). The GDPR requires that all EU citizens' data that are transferred to destinations outside the European Economic Area should be protected in a manner that is consistent with how personal data is protected in the EU.[69] Thus, the transmission may occur with countries that have an international agreement with the EU which follows the principles of international law, such as a mutual legal assistance treaty.[70] Where

no such decision from the EU Commission (EC) exists, transmission of personal data may be possible only subsequent to assessment of the adequacy of the level of protection by the EC. The Commission follows a set standard for assessment based on the rule of law, respect for human rights and fundamental freedoms, relevant legislations of such third country, and the effectiveness of the supervisory authorities and its commitments to protect personal data of individuals.[71]

However, if there is no adequacy decision for a country, this does not necessarily prevent any data transfer to said the third country. Under such circumstances, the controller must ensure that the recipient will protect the personal data. This can be assured using standard contractual clauses, binding corporate rules, and /or approved codes of conduct.[72] Furthermore, there could be data transferred to a third country even if the protection of personal data cannot be sufficiently guaranteed, provided the data subject has provided their consent.[73] This implies that where patient-related clinical data are being transmitted to third countries, it is not sufficient for informed consent for participation in the clinical study to exist. Rather, the informed consent must contain clear wording regarding transfer of study data to third countries or international organizations (as the case may be).

Such policies are increasingly being monitored to ensure enhanced levels of protection within a safer paradigm for data sharing. For instance, the US-EU policy ensuring safe harbor principles have been augmented with respect to data use and the onward transfer principle to offer privacy shield protection (Monteleone and Puccio 2017). Thus, data transfer to third parties must now also comply with the principle of purpose limitation of data and ensure that the third party provides the same level of protection in accordance with the so-called Privacy Shield Agreement.[74] The transparency of clinical data consequently comes with obligations for greater regulatory compliance particularly when transmitted internationally. Yet, legal uncertainty is persisting in this area due to recent US litigation (Minssen et al. 2019; Robins 2019) and since the current regulatory framework for international data transfer has been challenged again and is once more being considered by the CJEU (Baker 2019).[75]

6. Connecting the dots: GDPR, CTD transparency, and open innovation

As described above, the GDPR has now become the norm for the processing of personal data. It has increased data privacy, strengthened individual rights, and imposed specific obligations on the processing and governance of personal data. Whereas the governance of information exchange and data protection of traditional healthcare data and CTD has always been subject to very strict rules and controls, the full effect of the GDPR's broader application on more diverse data sets is now materializing. These developments have been considered essential in light of the recent Facebook and Cambridge Analytica scandals (Confessore 2018), which have drawn attention to the misuse of data. They have increased citizens' and patients' awareness of their data rights and have highlighted the fact that the use of healthcare records and CTD needs to be controlled and regulated.

At the same time, data transparency and different levels of data sharing have become increasingly important factors in the modern pharmaceutical innovation landscape demanding greater responsibility on the part of the stakeholders for sharing clean, well-described, and accurate data files that can be used by others without infringing upon privacy and patient confidentiality (Ross et al. 2012). There is no doubt that private actors and public regulators, such as the EMA, will increasingly refer to the massive amounts of data generated on a daily basis by wearable devices, electronic health records, social media, traditional clinical trials, or spontaneous adverse reaction report to assess the benefit–risk of medicines across their lifecycle.[76] Yet, the full benefits of such new and increasingly diverse big data and real-life evidence can only be realized if patient-level data are accessible and interoperable.

Within this trajectory, we identified various interrelated challenges at the interface between the obligations posed by the GDPR and the new realities for innovation and regulatory processes. In our view, these crystalize in the areas of (1) *transparency* (2) *open innovation*, (3) and *emerging technologies* to facilitate the use of big and FAIR (Wilkinson et al. 2016; Hodson et al. 2018)[77] data in clinical trials.

CTR and its transparency stipulations were adopted to support medical authorities and clinical trial sponsors by speeding up and simplifying large international clinical trials and the use of clinical data. Yet, the increasing significance and complexity of multi-data applications, their implications for the EMA's long term assessments of medicines, and their interplay with the GDPR had not fully been anticipated. In this article, we have identified and discussed four selected topical areas where this becomes evident in the context of clinical trials, namely informed consent, the use of public data, anonymization, and the regulatory regime for international data transfer, which is characterized by a persistent complexity and uncertainty. As we have demonstrated, within these areas it will be crucial to reconcile reasonable protection of privacy with the CTR's main objectives. It will, for example, be vital to reach an acceptable balance between effective data anonymization and maintenance of the data's clinical utility to the greatest possible extent.[78] In addition, controllers must be ready to look at all the circumstances involved and make reasonable and thorough determinations as to whether there is any imbalance of power prior to relying on consent as a legal basis. If the CTR's main objectives are unduly compromised by the fear of exposing personal data and the complexities of avoiding imbalance, this might not only undermine the rationales underpinning recent openness movements and clinical trials transparency regulations in CTR. It could also have a potentially negative effect on up-stream medical innovation and competition.

Turning to the *open innovation* aspects, we note that the future of clinical trials is aimed at using more and more complex data sets along with the integration of stakeholders other than the pharmaceutical industries themselves such as patients, doctors, and advocacy groups. This is essential in the interest of developing personalized treatments for rare diseases and new treatment protocols. In this sense, clinical trial research needs to be more open, nonlinear, and collaborative.[79] This implies the development and maintenance of wider pharmaceutical ecosystems in which a constellation of stakeholders will jointly be responsible for offering an integrative solution respecting the uncertainty and complexity of this context. This not only requires the development of mechanisms with which those stakeholders can collaborate in an open and inclusive way. It also demands strategies on how to cope with the dynamics that will shape the development of the ecosystem in terms of, for example, data access and appropriability, as well as the standardization of relevant interfaces (Holgersson et al. 2018).

Considering the longer-term viability of the pharmaceutical data ecosystem, there will be a trade-off with respect to the use of health

data which could limit the disruptive potential of open innovation efforts. Given the potentially broad territorial scope and the range of data sets, processes, and activities that are covered, it is evident that the GDPR will have a substantial impact on the generation, use, transparency, storage, and interoperability of health records. It will affect big companies, as well as small- and medium-sized enterprises (SMEs) (Price II et al. 2019). The industry must now be fully aware of their obligations as data processors and data controllers, which implies ensuring that every stage relevant to their activities recognizes the data flow and can show this audit trail to anyone, at any time (Price II et al. 2019). This also means that clinical trials and the authorities which regulate them, such as the EMA, need to follow—and potentially improve on—the standards that have been set in place in more traditional medical spheres (Price II et al. 2019). It will, therefore, be crucial for public policy and private organizations to develop compliance tools and discuss new technological solutions and amendments that would help start-ups, SMEs, and medical authorities to meet the compliance requirements without compromising too much of data's value. In summary, a more open approach to data sharing and transparency can only work if the entire ecosystem of complementary stakeholders manages to design compatible interfaces while respecting the regulatory framework, although this process will, in itself, be dynamic and iterative given the evolving nature of all these aspects.

This brings us to another issue: emerging technologies. Unfortunately, the GDPR may not be very well suited to fully accommodate the possibilities that are offered by new technologies which could be particularly useful for processing clinical trial data, such as blockchain technology (Siwicki 2018), which might help in addressing some tensions between data transparency and the requirements for personal data and IP protection. Clinical trial data represent a special data category and an appropriate consent can enable a research participant's rights to erasure, or portability, under the GDPR. However, where the consent is flawed, unclear, not robust or dynamic in nature, this could pose problems for new technological applications. As described by the head of Data Policy at the World Economic Forum, Anne Toth (Toth 2018):

> Because blockchain relies on a distributed ledger system that is decentralized and immutable, it's intended to be a permanent, tamper-proof record that sits outside the control of any one governing authority. This is what makes it such an attractive and useful technology. But because data that could be stored on the blockchain can't be deleted, there is no way to exercise the right to erasure that people are granted under GDPR. Without any further modifications Blockchain is not designed to be GDPR-compatible. Or rather, GDPR is not blockchain-compatible the way it is written today.

Similar problems may occur concerning the GDPR's definition of personal data. The GDPR's categorizations focus on the point in time when the data is being collected. Yet, in addition to assessing the appropriate anonymization technique, the data controller must also ensure that there is no risk of re-identification at later stages of cross-data use. These 'data dynamics' are not sufficiently addressed in the GDPR. In particular, when considering new technical developments with regard to data-triangulation or quantum computing, future changes to the legislation might be required to reconcile privacy concerns and societal control over new data and AI-driven technologies with the need for innovation- and competition-friendly regulatory framework.

Too many restrictions on the use of clinical trial data, however, might turn out to be an impediment to the full utilization of data and technological possibilities for the greatest public good and simultaneously downplay the benefits of openness. In addition, they could also be detrimental to competition in health-related technologies by favoring large players and impeding SMEs.[80]

Conceding these factual and legal concerns, it is clear that GDPR has a greater impact on data governance obligations for the sponsors, data controllers, and processors involved in the clinical trials. The GDPR also has an important impact on the legal basis for the processing of clinical data for primary and secondary uses and in the processing of the health data which are invariably involved in clinical trials. In addition, the GDPR has a greater impact on clinical research where the processing of data is for the purposes of scientific research involving technological development, fundamental or applied research, and privately funded research. Nevertheless, the GDPR does not offer the best legislative option for all the stakeholders involved in clinical trial research.

To address the limitations encountered by clinical trials in relation to the GDPR, legislators should monitor new technological developments very carefully and must be open to making amendments or adopting alternative measures wherever needed. In considering such changes, particularly in re-defining exemptions (Mostert et al. 2016), decision makers should be guided by three goals (Cohen and Michelle 2018):

1. Avoiding undue burdens on health research and public health activities,
2. Giving data subjects information on how their personal information is or could be used to the greatest extent commensurable with the first goal, and
3. Holding data users accountable for departures from authorized uses of data.

We believe that these high-level goals provide a way forward for realigning the clinical trial research within the general standards offered by the GDPR.

7. Concluding remarks

The interplay between new transparency policies and the GDPR results in various structural challenges and complications. This has substantial implications for how more openness can best be utilized to maximize public and private benefit. Clinical data transparency allows for the sharing of evidence-based medical information in the interest of different stakeholders such as patients, research, and the pharmaceuticals industry itself. At the same time, the GDPR ensures that individuals' rights and freedoms in relation to their personal data are sufficiently protected whenever various types of data, including research data, are being processed.[81] In this sense, both data protection under the GDPR and clinical data transparency under the CTR aim for the openness of information within set standards. The question is, in how far can the aims of both these regulations and policy initiatives be reconciled to meet their rationales and goals?; and related to that: What is their effect on data utility and innovation in the life sciences in particular?

The answer to the question rests partly on the outlook pursued in policy, and partly on the structural gaps which cannot enforce and reward adequate sharing of data at different levels in the context of openness. Sharing of information under the clinical transparency policies of the EMA was, until recently, limited to clinical

reports and did not offer information on the raw study data sets which are essential for evidence-based research and the verification of results.[82] Recent initiatives are changing this, but pharmaceutical companies will continue to exclude essential data under data protection rules, such as the GDPR's data minimization requirement.[83] Because drug development involves years of strategic investment and risk-taking in research and innovation, pharmaceutical companies also have a greater incentive to retain control over the data they harvest. Hence, claims of commercial confidentiality and the potential difficulties in de-identifying and sharing raw data might additionally impede progress in data transparency (Mintzes et al. 2015).

The conflicting obligations of privacy and transparency, as well as the motivation to protect valuable data through trade secrets and IPRs, could thus indeed undermine the potential benefits of open innovation and patient interests, whose high-quality health data is needed in order to develop innovative therapies more effectively (Lemmens 2013). The stakes are high since more extensive redactions of data might also compromise innovation, trust and data utility with regard to the detection of adverse effects. On the other hand, poorly administered wider transparency might also open doors for data misuses and piggy-backing on expensive research efforts, which could potentially have detrimental effects on trust and the incentives for companies to further investment in risky pharma projects.[84]

The work of the relevant data committees at the EMA in implementing these policies is therefore as important as it is complex. They must consider more sophisticated, yet still predictable and feasible measures to control, enhance and incentivize the sharing of data within an open innovation paradigm that can be reconciled with adequate protection of data and IPRs. This process requires coordination among multiple stakeholders ineffectively calling for a better alignment between research, practices, and policies (Bogers et al. 2018).

Acknowledgements

Special thanks go to CeBIL's student assistants Bénédicte Illien and Dimitrios Svolis for their valuable help in the preparation of this article.

Funding

Our research is supported by the scientifically independent Collaborative Research Program for Biomedical Innovation Law, which is supported by the Novo Nordisk Foundation (grant NNF17SA0027784). The funder had no role in the preparation, review, or approval of the manuscript and decision to submit the manuscript for publication.

Conflict of interest statement.None declared.

Notes

1. See EMA (2014) *EMA policy on publication of clinical data for medicinal products for human use*, POLICY/0070, EMA/240910/2013, (London, UK; pubd online Oct 2014) <http://www.ema.europa.eu/docs/en_GB/document_library/Other/2014/10/WC500174796.pdf> accessed 11 Aug 2019; cf. EMA (2018) *EMA's Clinical data publication (Policy 0070) report Oct 2016- Oct 2017*, EMA/630246/2017 (London, UK; pubd online July 2018), <http://www.ema.europa.eu/docs/en_GB/document_library/Report/2018/07/WC500252071.pdf> accessed 11 Aug 2019.
2. Regulation (EU) 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC (referred hereafter as 'Regulation (EU) 536/2014').
3. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation or GDPR).
4. SMA Europe, *Clinical Trials for SMA* <http://www.sma-europe.eu/clinical-trials/> accessed 11 December 2018. See also Martin et al. (2017).
5. An independent study helped the EMA to remove a dangerous drug from the European market in 2010, see: EMA (2010) *EMA recommends suspension of Avandia, Avandamet, and Avaglim* (London, UK; pubd online Oct 2010) <https://www.ema.europa.eu/en/news/european-medicines-agency-recommends-suspension-avandia-avandamet-avaglim> accessed 11 Aug 2019.
6. EMA (2018) *EMA policy on access to documents*, POLICY/0043, EMA/729522/2016 (London, UK; pubd online Oct 2018), <https://www.ema.europa.eu/documents/other/policy/0043-european-medicines-agency-policy-access-documents_en.pdf>.
7. EMA, Clinical data publication, <http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/general/general_content_000555.jsp&mid=WC0b01ac05809f363e> accessed 11 December 2018.
8. In AbbVie v EMA an interim relief against disclosure was sought, cf. CJEU, General Court, 23 April 2013, *AbbVie v EMA*, case T-44/13R. On appeal, this interim relief was set aside, cf. CJEU, Court of Justice, 28 November 2013, *EMA v AbbVie*, case C-389/13P.
9. Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information against their unlawful acquisition, use and disclosure, OJ L 157, 15.6.2016, p. 1–18.
10. See Case T-73/13, *InterMune UK a.o. v EMA* (29 June 2015); Case C-406/16 Pari Pharma v EMA, para 36 (18 October 2016); Case T-718/15, PTC Therapeutics International v EMA (5 February 2018); Case T 729/15, MSD Animal Health Innovation and Intervet international v EMA (5 February 2018); Case T-33/17, Amicus Therapeutics UK and Amicus Therapeutics v EMA (25 September 2018).
11. CJEU, Case T-235/15, *Pari Pharma v EMA*, para 36 (5 February 2018).
12. EMA (2014) *EMA policy on publication of clinical data for medicinal products for human use*, POLICY/0070, EMA/240910/2013, (London, UK; pubd online Oct 2014) <http://www.ema.europa.eu/docs/en_GB/document_library/Other/2014/10/WC500174796.pdf> accessed 11 Aug 2019; cf. EMA (2018) *EMA's Clinical data publication (Policy 0070) report Oct 2016- Oct 2017*, EMA/630246/2017 (London, UK; pubd online July 2018), <http://www.ema.europa.eu/docs/en_GB/document_library/Report/2018/07/WC500252071.pdf> accessed 11 Aug 2019.
13. Recital 67 Regulation (EU) 536/2014.
14. Article 81(4) Regulation (EU) 536/2014.
15. Regulation (EC) 1049/2001 of the European Parliament and of the Council of 30 May 2001 regarding public access to European Parliament, Council and Commission documents.

16. Article 81(7) Regulation (EU) 536/2014.
17. Article 81 (2) and (6) Regulation (EU) 536/2014.
18. EMA (2016) *External guidance on the implementation of the EMA policy on the publication of clinical data for medicinal products for human use*, EMA/90915/2016 (London, UK; pubd online Mar 2016) <https://www.ema.europa.eu/en/human-regulatory/marketing-authorization/clinical-data-publication/support-industry/external-guidance-implementation-european-medicines-agency-policy-publication-clinical-data> accessed 11 Aug 2019.
19. EMA, *Technical anonymization group* (London, UK) <http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_001880.jsp&mid=WC0b01ac0580c77e78> accessed 11 Aug 2019.
20. EMA (2018) *EMA's proactive publication of clinical data a success* (London, UK; pubd online July 2018) <https://www.ema.europa.eu/en/news/emas-proactive-publication-clinical-data-success> accessed 11 Aug 2019, stating: 'The report covers one year from the launch of EMA's clinical data website on 20 October 2016 and lists the 50 medicines for which clinical data were published, including orphan, pediatric, biosimilar and generic medicines, as well as the corresponding 54 regulatory dossiers. These data have attracted a total of 3,641 users, resulting in 22,164 document 'views' and 80,537 'downloads' for non-commercial research purposes. The report sheds light on the total number of documents published, the amount of commercially confidential information (CCI) redacted and the anonymization techniques used. EMA accepted 24% of CCI redactions proposed by pharmaceutical companies, with the result that only 0.01% of 1.3 million pages published contained CCI redactions.'
21. Id.
22. EMA, *Background to clinical data publication policy* (London, UK) <http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/general/general_content_000556.jsp&mid=WC0b01ac05809f363f> accessed 11 Aug 2019.
23. *Ibid*.
24. See Article 4 GDPR.
25. Health data is defined at Article 4(15) and recital 35 GDPR. Genetic data is defined at Article 4(13) and recital 34.
26. See Article 9, Article 4(13-15) and recitals 34 (on genetic data), 35, 45, 52-54, 63, 65, 71, 75, 91 GDPR.
27. Article 9 (g) and (i); recital 52.
28. Recitals 52 and 53.
29. See Article 9, Article 4(13-15) and recitals 34 (on genetic data), 35, 45, 52-54, 63, 65, 71, 75, 91.
30. Article 6 GDPR ; Article 1(c) GDPR (data minimization); Article 14 GDPR; Articles 24, 35 and 36; recitals 76 and 77; Article 25 GDPR. Id.
31. Article 40 GDPR.
32. Article 35 GDPR.
33. Recital 47 GDPR.
34. European Data Protection Board (2019) *Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR)* (pubd online Jan 2019) <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_opinionctrq_a_final_en.pdf> accessed 11 Aug 2019.
35. Article 6(1) (c), Article 9(2)(i) GDPR.
36. Organization for Economic Co-operation and Development, The OECD Privacy Framework (2013) 'Guidelines Governing the Protection of Privacy and Trans-border Data Flows of Personal Data', <http://www.oecd.org/internet/ieconomy/privacy-guidelines.htm> accessed 11 December 2018; see also 4th International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH Harmonized Tripartite Guideline (1996) *Guideline for Good Clinical Practice E6 (R1)* <http://apps.who.int/medicinedocs/documents/s22154en/s22154en.pdf>.
37. Recital 27 Regulation (EU) 536/2014.
38. Article 28 (b), (c) Regulation (EU) 536/2014.
39. Article 10 Regulation (EU) 536/2014.
40. Article 28 (2) Regulation (EU) 536/2014.
41. Articles 6, 7, 9 GDPR.
42. Article 9 GDPR.
43. European Data Protection Supervisor (2013) *Guidelines on the processing of personal data in the context of public procurement, grants as well as selection and use of external experts* (pubd online June 2013) <https://edps.europa.eu/sites/edp/files/publication/13-06-25_procurement_en.pdf> accessed 11 Aug 2019.
44. Article 2(21) Regulation (EU) 536/2014.
45. Article 89, Article 21 (6), Article 5(e) GDPR.
46. European Commission, Guidelines on Consent under Regulation 2016/679, 6 July 2018, <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051> accessed 11 December 2018.
47. Recital 43 GDPR.
48. Article 4 and Recital 43 GDPR.
49. See also European Patient Forum (2016) *Clinical Trials Regulation: Informed Consent and Information to Patients* (pubd online May 2016) <http://www.eu-patient.eu/globalassets/policy/clinicaltrials/epf_informed_consent_position_statement_may16.pdf> accessed 11 Aug 2019.
50. EMA, *supra* note 1.
51. Recital 26 GDPR stating that the GDPR may apply to pseudonymized but generally not to anonymized data. Pseudonymization is a data management procedure by which personally identifiable information fields within a consumer data record are replaced by one or more artificial identifiers, or pseudonyms, which may be recalled at a later date to re-identify the record. In contrast, anonymization is the process of either encrypting or removing personally identifiable information from data sets so that the people whom the data describes remain permanently anonymous.
52. Article 3 GDPR.
53. The Article 29 Working Party uses the combined expression 'inferred data and derived data', see European Commission (2017) *Guidelines on the right to 'data portability'* (pubd online Oct 2017), <http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233> accessed 11 Aug 2019.
54. *Ibid*.
55. EMA, *supra* note 1.
56. CJEU, 17 July 2014, *YS v Minister voor Immigratie, Integratie en Asiel and Minister voor Immigratie, Integratie en Asiel v M and S*, case C-141/12.
57. *Ibid*. para 33-48.
58. EMA, *Clinical data publication, Terms of use* (London, UK) <https://clinicaldata.ema.europa.eu/web/cdp/termsofuse> accessed 11 Dec 2018

59. EMA, *supra* note 18.
60. Ibid.
61. Biobanking and Biomolecular Resources Research Infrastructure (2016) *The EU General Data Protection Regulation – Answers to the frequently asked questions* <http://www.bbmri-eric.eu/wp-content/uploads/2016/08/BBMRI-ERIC_FAQs_on_the_GDPR_V1.0_for_print.pdf> accessed 11 Dec 2018.
62. Recital 26 GDPR.
63. EMA, *supra* note 40.
64. International Organization for Standardization (2008) *ISO/TS 25237:2008 - Health informatics – Pseudonymization* (pubd online Dec 2008) <https://www.iso.org/standard/42807.html> accessed 11 Aug 2019; Article 29 Working Party (2014) *Opinion 05/2014 on Anonymization Techniques*, 0829/14/EN WP216.
65. *Ibid*.
66. Article 29 Working Party, *supra* note 53.
67. EMA, *supra* note 18.
68. El Emam K., and Aballah K. (2015) 'De-identifying Clinical Trials Data', Applied Clinical Trials http://www.appliedclinicaltrialsonline.com/de-identifying-clinical-trials-data accessed 11 Dec 2018.
69. Recital 101 and Article 44 GDPR.
70. Recital 115 GDPR.
71. Article 45 GDPR; See also adequacy decisions at <https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en>
72. Articles 42, 43, 48, 49 GDPR; see also CJEU, case C-362/14, Maximillian Schrems v Data Protection Commissioner (6 October 2015).
73. Article 49 GDPR.
74. See European Commission (2016) *EU Commission and United States agree on new framework for transatlantic data flows: EU-US Privacy Shield* (pubd online Feb 2016) <http://europa.eu/rapid/press-release_IP-16-216_en.htm> accessed 29 June 2019.
75. See Case T-738/16, *Quadrature du Net v. Commission*, where hearings have been suspended due to the pending resolution of case C-311/18 Facebook Ireland & Schrems (hearing due on 9th July).
76. HMA-EMA Joint Big Data Taskforce, *Summary report* (online pub Feb 2019) <https://www.ema.europa.eu/documents/minutes/hma/ema-joint-task-force-big-data-summary-report_en.pdf> accessed 11 Aug 2019.
77. FAIR data = (F)indable, (A)ccesible, (I)nteroperable and (R)eusable data.
78. EMA (2018) *Data anonymization - a key enabler for clinical data sharing*, Workshop report, EMA/796532/2018 (London, UK; pubd online Dec 2018) <https://mrctcenter.org/wp-content/uploads/2018/12/Data-anonymization-a-key-enabler-for-clinical-data-sharing-Workshop-pdf> accessed 11 Aug 2019.
79. See for instance Integrated Medicines Initiative (2017) *Personalized medicine through open innovation* (pubd online Nov 2017) <https://www.imi.europa.eu/news-events/events/personalized-medicine-through-open-innovation> accessed 11 Aug 2019.
80. Further research on the interplay between IPR, data protection rules and competition law is crucial. However, this was not at the focus of this paper.
81. Recital 46 47, GDPR.
82. See for instance: House of Commons, Science and Technology Committee (2013) *Third report, Clinical trials*, (pubd online Sept 2013) <https://publications.parliament.uk/pa/cm201314/cmselect/cmsctech/104/10402.htm> accessed 11 Aug 2019.
83. Article 5 GDPR: 'Personal data shall be: (...) (c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ("data minimization")'.
84. Pharmaceutical Research and Manufacturers of America, European Federation of Pharmaceutical Industries and Associations (2013) *Principles for responsible clinical trial data sharing: our commitment to patients and researchers* (pubd online July 2013) <http://phrma-docs.phrma.org/sites/default/files/pdf/PhRMAPrinciplesForResponsibleClinicalTrialDataSharing.pdf> accessed 11 Aug 2019.

References

Baker, J. (2019), *EU High Court Hearings to Determine Future of Privacy Shield, SCCs*, (Portsmouth, NH; pubd online June 2019) <https://iapp.org/news/a/eu-high-court-hearings-to-determine-future-of-privacy-shield-standard-contractual-clauses/> accessed 11 Aug 2019.

Bhardwaj, A., Scaria V., Raghava G. P. et al. (2011) 'Open Source Drug Discovery: A New Paradigm of Collaborative Research in Tuberculosis Drug Development', *Tuberculosis*, 91/5: 479–86.

Bogers, M. (2011) 'The Open Innovation Paradox: Knowledge Sharing and Protection in R&D Collaborations', *European Journal of Innovation Management*, 14/1: 93–117.

———, Chesbrough, H., and Moedas, C. (2018) 'Open Innovation: Research, Practices, and Policies', *California Management Review*, 60/2: 5–16.

Bonini, S., Eichler H. G., Wathion N. et al. (2014) 'Transparency and the EMA-Sharing of Clinical Trial Data', *The New England Journal of Medicine*, 371: 2452–5.

Chesbrough, H., and Bogers, M. (2014) 'Explicating Open Innovation: Clarifying an Emerging Paradigm for Understanding Innovation'. In: H., Chesbrough, W., Vanhaverbeke, and J., West (eds.) *New Frontiers in Open Innovation*, pp. 3–28. Oxford: Oxford University Press.

Cohen, I. G., and Michelle, M. M. (2018) 'HIPAA and Protecting Health Information in the 21st Century', *Journal of the American Medical Association*, 320/3: 231–2.

———, and Mello, M. M. (2019) 'Big Data, Big Tech, and Protecting Patient Privacy', *Journal of the American Medical Association*, 322/12: 1141–2.

Confessore, N. (2018) *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far* (pubd online April 2018) <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scand al-fallout.html> accessed 11 Aug 2019.

Dias, M. (2015) 'Guidance on the Anonymization of Clinical Reports for the Purpose of Publication in Accordance with Policy 0070'. Presented at the EMA, Amsterdam, Netherland s, 6 July 2015.

Donovan, F. (2018) *Does Quantum Computing Threaten Healthcare Data Security?* (pubd online April 2018) <https://healthitsecurity.com/news/does-quantum-computing-threaten-healthcare-data-security> accessed 11 Aug 2019.

El Emam, K., and Aballah, K. (2015) *De-identifying Clinical Trials Data* (pubd online March 2015) <http://www.appliedclinicaltrialsonline.com/de-identifying-clinical-trials-data> accessed 11 Aug 2019.

Gogates, G. (2018) *How Does GDPR Affect Clinical Trials?* (pubd online April 2018) <http://www.appliedclinicaltrialsonline.com/how-does-gdpr-affect-clinical-trials> accessed 11 Aug 2019.

Harford, K., Mulryne, J., Williams, A. et al. (2018) *Clinical Trial Consents Under the EU GDPR: Where Do We Stand?* (pubd online June 2018) <https://www.biosliceblog.com/2018/06/clinical-trial-consents-eu-gdpr-stand /> accessed 11 Aug 2019.

Hodson, S., Jones, S., Collins, S. et al. (2018) 'Turning FAIR Data into Reality: Interim Report from the European Commission Expert Group on FAIR Data', Zenodo. doi : 10.5281/ZENODO.1285272.

Holgersson, M., Granstrand, O., and Bogers, M. (2018) 'The Evolution of Intellectual Property Strategy in Innovation Ecosystems: Uncovering Complementary and Substitute Appropriability Regimes', *Long Range Planning*, 51/2: 303–19.

Hudson, K. L., Lauer, M. S., and Collins, F. S. (2016) 'Toward a New Era of Trust and Transparency in Clinical Trials', *JAMA*, 316/13: 1353–4.

Hughes, S., Wells, K. McSorley, P. et al. (2014) 'Preparing Individual Patient Data from Clinical Trials for Sharing: The GlaxoSmithKline Approach', *Pharmaceutical Statistics*, 13/3: 179–83.

Institute of Medicine. (2015) *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. Washington, DC: The National Academies Press.

Kaye, J., Whitley, E. A., Lund, D. et al. (2015) 'Dynamic Consent: A Patient Interface for Twenty-First Century Research Networks', *European Journal of Human Genetics*, 23: 141–6.

Kmietowicz, Z. (2014) 'Transparency Campaigners Welcome New Rules for Clinical Trials in Europe', *BMJ*, 348: g2579.

Lassman, S. M., Shopshear, O. M., Jazic, I. et al. (2017) 'Clinical Trial Transparency: A Reassessment of Industry Compliance with Clinical Trial Registration and Reporting Requirements in the United States', *BMJ Open*, 7: e015110.

Lemmens, T. (2013) 'Pharmaceutical Knowledge Governance: A Human Rights Perspective', *Journal of Law Medicines and Ethics*, 41/1: 163–84.

Lexchin, J., Bero, L., Djulbegovic, B., and Clark, O. (2003) 'Pharmaceutical Industry Sponsorship and Research Outcome and Quality: Systematic Review', *British Medical Journal*, 326: 1167–70.

Longo, D. L., and Drazen, J. M. (2016) 'Data Sharing', *The New England Journal of Medicine*, 374/3: 276–7.

Lowman, M., Trott, P., Hoecht, A. et al. (2012) 'Innovation Risks of Outsourcing in Pharmaceutical New Product Development', *Technovation*, 32/2: 99–109.

Martin, B. R. (2016) 'Twenty Challenges for Innovation Studies', *Science and Public Policy*, 43/3: 432–50.

Martin, L., Hutchens, M., Hawkins, C. et al. (2017) 'How Much Do Clinical Trials Cost', *Nature Reviews Drug Discovery*, 16: 381–2.

Mello, M. M., Francer, J. K., Wilenzick, M. et al. (2013) 'Preparing for Responsible Sharing of Clinical Trial Data', *The New England Journal of Medicine*, 369/17: 1651–8.

Mintzes, B., Lexchin, J., and la Santos Quintano, A. (2015) 'Clinical Trial Transparency: Many Gains But Access to Evidence for New Medicines Remains Imperfect', *British Medical Bulletin*, 116/1: 43–53.

Minssen, T., Gerke, S., and Shachar, C. (2019) *Is Data Sharing Caring Enough About Patient Privacy? Part I and Part II* (pubd online July 2019) <http://blog.petrieflom.law.harvard.edu/2019/07/26/is-data-sharing-caring-enough-about-patient-privacy-part-i-the-background/> accessed 2 Sept 2019.

Mohs, R. C., and Greig, N. H. (2017) 'Drug Discovery and Development: Role of Basic Biological Research', *Alzheimers Dement*, 3/4: 651–7.

Monteleone, S., and Puccio, L. (2017) *From Safe Harbour to Privacy Shield: Advances and Shortcomings of the New EU-US Data Transfer Rules* (pubd online Jan 2017) <http://www.europarl.europa.eu/thinktank/en/document.html? reference=EPRS_IDA(2017)595892> accessed 11 Aug 2019.

Mostert, M., Bredenoord, A. L, Biesaart, M. C. et al. (2016) 'Big Data in Medical Research and EU Data Protection Law: Challenges to the Consent or Anonymise Approach', *European Journal of Human Genetics*, 24/7: 956–60.

Packer, M. (2018) 'Data Sharing in Medical Research', *BMJ*, 360: k510.

PhUSE De-Identification Working Group. (2015) *Providing De-Identification Standards for CDISC Data Models* (pubd online 2015) <https://pharmasug.org/proceedings/2015/DS/PharmaSUG-2015-DS10.pdf> accessed 11 Aug 2019.

Price, W. N., II., and Minssen, T. (2015) 'Will Clinical Trial Data Disclosure Reduce Incentives to Develop New Uses of Drugs', *Nature Biotechnology*, 33: 685–6.

———, Kaminski, M. E., Minssen, T. et al. (2019) 'Shadow Health Records Meet New Data Privacy Laws: How Will Research Respond to a Changing Regulatory Space', *Science*, 363/6426: 449.

———., and Cohen, I. G. (2019) 'Privacy in the Age of Medical Big Data', *Nature Medicine*, 25: 37–43.

Rajam, N., and Minssen, T. (2019) 'What Lurks in the Shadows of the Openness Hyperbole for Biopharmaceuticals', *Drug Development Research*, 80/3: 282.

Robins, R. (2019) *Potential Class Action Lawsuit Accuses the University of Chicago of Sharing Identifiable Patient Data with Google* (pubd online June 2019) <https://www.statnews.com/2019/06/26/potential-class-action-lawsuit-accuses-the-university-of-chicago-of-sharing-identifiable-patient-data-with-google/> accessed 29 Jun 2019.

Ross, J. S., Lehman, R., Gross, C. P. et al. (2012) 'The Importance of Clinical Trial Data Sharing: Toward More Open Science', *Circulation, Cardiovascular Quality and Outcomes*, 5/2: 238–40.

Schuhmacher, A., Gassmann, O., and Hinder, M. (2016) 'Changing R&D Models in Research-Based Pharmaceutical Companies', *Journal of Translational Medicine*, 14/1: 105.

Siwicki, B. (2018) *How One Startup is Bringing Block Chain to Genomics and Precision Medicine* (pubd online July 2018) <https://www.healthcareitnews.com/news/how-one-startup-bringing-blockchain-genomics-and-precision-medicine>.

SMA Europe. (2019) *Clinical Trials for SMA* <http://www.sma-europe.eu/clinical-trials/> accessed 11 Dec 2019.

Toga, A. W., Foster, I., Kesselman, C. et al. (2015) 'Big Biomedical Data as the Key Resource for Discovery Science', *Journal of the American Medical Informatics Association*, 22/6: 1126–31.

Toth, A. (2018) *Will GDPR Block Block Chain?* (pubd online May 2018) <https://www.weforum.org/agenda/2018/05/will-gdpr-block-blockchain/> accessed 11 Aug 2019.

Travis, K. (2011) *Sharing Data in Biomedical and Clinical Research* (pubd online February 2011) <http://dx.doi.org/10.1126/science.caredit.a1100014> accessed 11 Aug 2019.

Tucker, K., Branson, J., Dilleen, M. et al. (2016) 'Protecting Patient Privacy When Sharing Patient-Level Data from Clinical Trials', *BMC Medical Research Methodology*, 16, 77.

Van Der Marel, E., Lee-Makiyama, H., and Bauer, M. (2014) *The Cost of Data Localisation: Friendly Fire on Economic Recovery* (pubd online May 2014) <http://ecipe.org/publications/dataloc/> accessed 11 Aug 2019.

Watson, R. (2014) 'European Medicines Agency Changes Policy on Clinical Trial Data Publication', *BMJ Clinical Research*, 4073: g4073.

Wilkinson, M. D., Dumontier M, Aalbersberg, I. J. et al. (2016) 'The FAIR Guiding Principles for Scientific Data Management and Stewardship', *Scientific Data*, 3: 160018. doi:10.1038/sdata.2016.18.