International Sharing of Personal Health Data for Research

April 2021

allea | All European Academies

European Academies
easac
Science Advisory Council

FEAM
Federation of European Academies of Medicine

Date of publication: 08 April 2021

DOI: www.doi.org/10.26356/IHDT

ISBN 978-3-8047-4249-9

Cover picture is by Shutterstock.

International Sharing of Personal Health Data for Research

April 2021

The ALLEA, EASAC and FEAM joint initiative on resolving the barriers of transferring public sector data outside the EU/EEA

Contents

Foreword

Health research is crucial for all. It benefits individual patients and populations, supports development of health care systems, and underpins social cohesion and stability. Collecting and combining health data is fundamental for the advancement of medical research, reducing health inequalities, and improving disease diagnosis and treatment. Sharing pseudonymised personal health data for public sector research is also essential to make most effective use of limited resources.

Sharing data safely and effectively must take account of privacy concerns. However, it has become apparent that implementation of the EU's General Data Protection Regulation (GDPR) has contributed to barriers in sharing health data with researchers outside the EU/EEA. A major challenge is the statutory conflict between EU fundamental rights and other countries' legislation. This obstacle affects the transfer of data to foreign institutions and also remote access by other researchers to data at its original location. Both activities are essential for international collaborative research. When institutions in other countries have statutory conflicts that prevent them from signing the required contracts under the GDPR, there is currently no workable legal mechanism for sharing health data outside the EU/EEA for public sector research.

The EU/EEA has had a great history of collaborative health research and has been a world leader in many of the areas of critical importance for addressing societal priorities. This leadership position is now at risk. While

discussion elsewhere of problems associated with implementation of the GDPR has mostly focused on the sharing of data within the private sector, the problems for public sector researchers have been neglected by policy makers. This neglect must be corrected rapidly because the problems affect patients and all citizens who are beneficiaries of public sector health research.

In view of the great importance of these issues, the European academy networks, ALLEA, EASAC and FEAM came together for their first tripartite project and this report is the result. Our main objectives are first, to emphasise the vital importance of the value of health research that is now in danger of being lost and second, to offer guidance to resolve the escalating problem, while respecting the right to protection of personal data. There is pressing need to find a simple solution that is safe and respectful of fundamental rights and does not conflict with other countries' laws or with the regulations of international organisations. Our consensus report provides detailed analysis of the current situation and suggests options for reform. The EU cannot act alone and we also advise on the need for the EU to lead international discussion on agreeing principles and on action to remedy the problems.

The report has been prepared by consultation with a group of experts nominated by their national academies. We thank them and we also thank the independent peer reviewers, and the academies of ALLEA, EASAC and FEAM for their guidance and for their continuing commitment to communicate our shared messages at the national level as well as to the EU Institutions.

We welcome discussion of any of the points raised in our report or on related issues that merit attention. Less global sharing of health data for research is hurting everyone and we need to act urgently.

– Prof. George Griffin, FEAM President
– Prof. Antonio Loprieno, ALLEA President
– Prof. Christina Moberg, EASAC President

Summary

Personal health data provide a vital resource for research to save and improve lives, reduce health inequalities and benefit society. Research data should be regarded as a global public good. Sharing of data, including genetic and other health-related data, is an essential part of public sector medical research for improved health care and disease prevention, for example to ensure sufficiently large sample sizes, identify complex pathways, and compare the determinants and outcomes of disease in different settings, thereby making the most of the contribution by patients and volunteers to research. It is important for EU citizens that their data are shared for health research, to ascertain whether research results from elsewhere are relevant to their particular genetic makeup and risk factors.

At the same time, it is essential to provide appropriate protections for personal data privacy. The General Data Protection Regulation (GDPR) addresses the protection of personal data in the European Union (EU) and European Economic Area (EEA) and the international transfer of data to areas outside the region. It has become apparent that the implementation of the GDPR has introduced impediments to this international transfer of data to outside the EU/EEA, creating problems for academic researchers, health-care professionals and others in the public sector. These problems affect patients and all citizens who are the ultimate beneficiaries of public sector health research.

The present report is produced by an initiative of the European academy networks, ALLEA (the European Federation of Academies of Sciences and Humanities), EASAC (the European Academies' Science Advisory Council) and FEAM (the Federation of European Academies of Medicine), to reaffirm the vital importance of sharing personal health data for research in the public sector, to explore the issues for international transfer and to offer guidance to resolve the growing problem, while respecting the right to protection of personal data. Commentaries elsewhere on controversies surrounding the transfer of data under the GDPR have usually focused on the private sector and the voice of the public sector researcher has been relatively neglected. This must change.

The transfer of personal data for research outside the EU/EEA is a particular problem. Drawing on the Working Group discussions and other material described in our report, ALLEA, EASAC and FEAM have developed consensus messages that can be summarised as follows.

- **Health research is crucial for all:** for patient benefit, population health, development of health-care systems, and for social cohesion and stability.

- **Sharing pseudonymised personal health data for public sector research is essential:** strong pseudonymisation procedures are important to make effective use of limited resources and maximise the value of contributions made to research by patients and volunteers. Long-term structured anonymised health data are also important for further development of new areas of

research such as artificial intelligence.

• **Data must be shared safely and efficiently, taking account of privacy concerns:** this is part of the conduct of responsible science, and addressing these opportunities should be part of wider initiatives to build trust in research and researchers and to take account of patient views.

• **Implementation of the GDPR has resulted in impediments to data sharing with researchers outside the EU/EEA:** this affects both the direct transfer of data and remote access to data at its original location and secondary uses of the data by foreign institutions, all of which often representing collaborative research with EU researchers. Thus, European researchers and EU citizens benefit from international sharing of data. When other countries do not have equivalent procedures for data protection (adequacy) there is currently no workable mechanism for sharing health data for public sector research. It has been estimated that in 2019 more than 5,000 collaborative projects (projects involving the US National Institutes of Health and EEA countries) were affected[1], a solution is urgently needed both for ongoing collaborations as well as for new research.

• **There must be increased commitment to finding a solution to overcome the barriers in sharing data:** the preferred option is to find a solution under Article 46 of the GDPR with additional operational guidance provided by the European Data Protection Board accompanied by tangible examples to show how to apply the guidance to health research. This is urgent.

• **There must also be increased commitment to enabling the use of shareable data:** even when appropriate mechanisms for transferring data are established, there are other, methodological and technical quality issues that need resolving to enable interoperability in the use of data. These challenges require greater attention across the research community. Along with data, biological samples must sometimes be shipped to laboratories outside the EEA and temporary solutions (such as sending syntaxes) do not allow for highly specialised analyses—any legal mechanism for sharing data must also allow for analyses of such samples.

• **Privacy-enhancing technologies are relevant** in offering potential to improve data security but their use does not circumvent the requirements of the GDPR, nor do these technologies solve the problems presented here.

Continuing monitoring and assessment of the issues is imperative because of the fast-changing environment and technology development, other country initiatives on data sharing, the momentum favouring open science and data, the role of big data and artificial intelligence on large data analysis, and new opportunities and needs in health care and disease prevention. We recommend development of an interdisciplinary mechanism for continuing monitoring of these developments, acting also to raise visibility of the opportunities and challenges for personal health data sharing for research. We suggest that academies and their networks might adopt a primary role in catalysing discussion with

1 | http://www.iscintelligence.com/archivos_subidos/robert_eiss_gdpr_us-eu_cooperation_in_biomedical_science_isc_gdpr_seminar_19_nov_2019.pdf

research institutions and research funders, and in engagement with other stakeholders, particularly patient groups and policy-makers.

There is an indispensable role for the EU to lead global discussion about privacy rules, the value of health research and the free movement of data including options for reforming regulations in other countries for reciprocity in data sharing.

Our messages are addressed to Directorates across the European Commission, particularly those responsible for Justice, Health and Research, the European Data Protection Board, European Data Protection Supervisor, other EU institutions, policy-makers in Member States, and to all those in the scientific, medical and policy communities worldwide who are interested in, and affected by, these issues. We welcome the commitment by the European Commission, following previous inputs from the research community, to develop consistency across the GDPR and facilitation of cross-border sharing of personal data to support health research; but given the large number of ongoing collaborations between European researchers and institutions outside the EU/EEA on critical health research, we emphasise the urgency in tackling the problems.

During a late stage in our Working Group discussion, the European Data Protection Board (EDPB) and the European Commission published new proposals for certain key issues (in particular the EDPB roadmap, supplementary measures and standard contractual clauses (SCCs)) and these will be discussed subsequently. However, these recent recommendations from the EDPB and the European Commission do not solve the problems for the research community in transferring data outside the EU/EEA, nor are there transfer mechanisms that enable European researchers to place data into large databases that applicants can access. These problems must be solved very soon if European research potential is to be realised and if further research disconnects are to be obviated.

Less global sharing of health data for research is hurting everyone. The immediate challenge is to find a simple solution that is safe and respectful of fundamental rights, including the right to data protection and the right to effective remedy, and one that does not conflict with other countries' laws or with regulations of international organisations.

1 | Introduction: personal data and health

The General Data Protection Regulation (GDPR, 2016/679) addresses the protection of personal data in the European Union (EU) and European Economic Area (EEA), and the international transfer of personal data[2] outside these areas to "third" (non-EEA) countries and international organisations. Although most of the principles embedded in the GDPR are not new and the focus of protections has been primarily on the corporate sector, and although the GDPR acknowledges the importance of research, it has become apparent that the implementation of GDPR restrictions has created new impediments for academic researchers, health-care professionals and others in the public sector. This problem affects patients and citizens, who are the ultimate beneficiaries of health research.

The present report is produced by an initiative of the European academy networks, ALLEA, EASAC and FEAM, to reaffirm the vital importance of sharing pseudonymised (see the Glossary on page 53) health data for research within the public sector, to explore the issues for international transfer under the provisions of the GDPR and to provide guidance on how to resolve this growing problem while respecting the right to protection of personal data, which, together with securing the free movement of personal data, are the main objectives of the GDPR.

1.1 | Sharing matters: research data as a global public good

Personal data provide a vital resource for health research, improving consistency and validation to save and improve the lives of patients, reduce health inequalities and benefit society. While there may be many steps involved between conducting research and benefiting patients and society, sharing data when appropriate is a necessary part of fostering and translating research into practice for improved health care and disease prevention. Many agree that, when published, research data should be accessible to other researchers for specific, well-defined purposes but there are also significant opportunities for the sharing of data as part of the research process itself.

2| See the Glossary on page 50 and Box 1 for definitions.

Box 1 | Personal data, anonymisation and pseudonymisation

Information that can be traced back to an identified or identifiable individual constitutes personal data. Anonymous data, on the other hand, does not fall within the scope of the GDPR (only personal data, or those related to an identified or identifiable natural person are covered by the legislation). Anonymisation (where no keys to original identifiers are retained) consists of the removal of all attributes associated with an individual that could make them identifiable in a dataset. Rendering data anonymous would therefore cause such data to be out of the scope of the GDPR. If the data are uniquely identifiable (for instance a genome sequence) or the data are sufficiently rich to make it possible to identify an individual either on the basis of the data alone or by linkage with other data sources, the data are not considered anonymous. At the same time, rendering health data anonymous could potentially diminish its usefulness for research; for instance, anonymisation of data in epidemiological studies may require several steps where variables are categorised or eliminated, which may lead to diminished value of such data (Shabani and Borry 2018; Mascalzoni et al. 2019).

While anonymisation renders data non-identifiable (Shabani and Borry 2018), other de-identification techniques may be used to supplement privacy protection (Kaissis et al. 2020; see also section 4.5 and Appendix 3).

Different from this, in pseudonymisation, identifiable information fields are replaced by artificial codes or identifiers. A key is kept, linking the identifiable information with the code used in the data set. The key is locked up separately from the other data. Pseudonymisation does not diminish the usefulness of data for research, and it is the standard operating procedure for most scientific studies. Throughout this document, we refer to the sharing of pseudonymised data for research purposes.

See also the Glossary for definitions.

Research data are a global public good (Knottnerus 2015)[3] and most would agree that secure access – for specific purposes – to research data already collected should be allowed. Sharing personal data for health research within the public sector can bring future benefit to individual patients and to population health, and to European society as a whole through promoting social cohesion. There are strong ethical arguments in favour of sharing of data to make the most of research and the patient's/volunteer's contribution to research. Examples of the European value added previously by sharing personal health data for research include multi-national studies on risk factors for blood pressure and for suicide in schizophrenia, exploration of the link between diabetes treatment and the occurrence of cancer, and demonstration of

3| For further discussion of developing principles in regarding knowledge as a shared resource and a public good, and how to define, protect and build the knowledge commons in the digital age, see Hess and Ostrom (2011).

the association between smoking and lung cancer. Sharing data is vital throughout health research; some current priorities for sharing data discussed by the Working Group include studies of cancer heterogeneity, antimicrobial resistance, psychiatric disorders, genetic research in various clinical indications, and vaccine research.[4] There is added value in international data sharing with countries outside the EEA/EU when it is necessary, for example to compare a wider range of socio-economic determinants and genetic factors in disease, where disease subtypes may be comparatively rare and to compare different approaches to prevention and treatment. It is essential to collaborate by sharing samples and data from EU citizens to ensure that conclusions from international studies are valid for EU populations with their particular genetics, risk factors, and other environmental and social determinants of health.

By mobilising the research community to maximise the public benefit of the rapidly increasing amount of data, wider sharing can accelerate the pace of discovery and help increase efficiency and efficacy in data analysis, validation and utilisation. Data sharing is an essential part of modern research and, within medical research, data on individuals are often pooled to ensure sufficiently large study sample size, and to replicate findings and identify complex pathways. For patients with rare diseases, the sharing of data collected decades ago on an international basis may be essential because of the small number of patients present in any single country. There are also unprecedented opportunities for real-time diagnostic and therapeutic decision-making if significant amounts of data are shared. Data sharing has been routine in some research areas, for example genetics and genomics, but until relatively recently it has perhaps been less common within the public health research community (Walport and Brest 2011), although the benefits of registry-based research, for example, are acknowledged in the GDPR. Individual patient records can form the basis of observational studies of natural courses of the factors influencing health and disease, and help researchers identify suitable patients to participate in clinical trials (Fears et al. 2014). By re-using patient research data where appropriate, participants in clinical trials are then assured that the data they contribute help to further knowledge without unnecessary duplication of studies. *"Health and disease are global… Any research and action plan must include a European- and world-wide dialogue"* (EU Scientific Panel for Health 2016). EU researchers and EU citizens both gain from wider collaborative research.

Because of the obligation to optimise the potential gains for patients in sharing data for research, it is essential to do this while protecting citizens' privacy and promoting patient's trust in research. Countries also have strong obligations to protect and promote the health of their citizens. Contemporary science requires large international collaborations and, to a certain extent, clinical practice requires the comparison of data coming from different patients for best practices. This creates strong ethical obligations to share data as well as to ensure that data sharing is done in a way that fosters high standards of public trust and confidence. The human right to health has also been increasingly interpreted as including the human right to benefit from health research (Knoppers 2018).

4| Examples are discussed in further detail in chapter 2 and Appendix 2.

The evidence for, and the determinants of, public willingness to share their data for health research requires further assessment, although the perspective, for example from the European Patients' Forum, is that most patients would like their data to be used in public sector research. A systematic review of the literature on public attitudes (Kalkman et al. 2019a) found widespread public support for health data sharing for research, but this support is conditional on a governance framework that incorporates patients' values and needs, particularly for confidentiality. A major European survey on patients with rare diseases (see Courbier et al. 2019) also found support for data sharing to foster research and improve health care but, again, dependent on specific requirements, for example to respect privacy. However, a recent large survey of public perceptions in 22 countries (Middleton et al. 2020) found that willingness to donate one's own data for research is relatively low, particularly if there were to be multiple users of the data. This study concluded that more needs to be done to show that the research community is worthy of public trust and that data sharing is integral to making the most of the research undertaken. More needs to be done to take account of the views of patients.

1.2 | Lessons from COVID-19

The value of recent initiatives on COVID-19 patient data for research, to share research outputs immediately and to accelerate collaboration in research, has been widely acknowledged (see, for example, Moorthy et al. 2020; COVID-19 Clinical Research Coalition 2020). Guidance on data protection during the initial phase of the COVID-19 pandemic has been published by the European Data Protection Board (EDPB, Guidelines 03/20 adopted on 21 April 2020) and there will need to be continuing reflection on the lessons learned for rapid and timely data sharing during COVID-19. The pandemic might be regarded as an exceptional example for mandating research coordination because of overriding public interest. However, this experience has re-emphasised the necessity of international collaboration in research for both communicable and non-communicable diseases—and should also act as a stimulus for the research and policy communities to re-examine whether current standard procedures of data governance are adequate for making the most of personal data for public health research and action in a newly uncertain and rapidly changing world.

1.3 | Protecting patients and data while promoting research

Making personal data[5] accessible for research requires careful consideration of some critical management issues by the research community, in particular for preserving confidentiality and societal values, ensuring data quality and validity standards and interoperability, and re-examining the competitive system of rewards for research achievement (Knottnerus 2015; Ohmann et al. 2017). The term "shareable landscape"

5| Further discussion of the nature of identifiable data and its handling within the EU is provided, for example, by the European Patients' Forum "The new EU Regulation on the protection of personal data: what does it mean for patients?" https://www.eu-patient.eu/globalassets/policy/data-protection/data-protection-guide-for-patients-organisations.pdf. There are, however, different types of health data which are collected under different conditions in different countries and according to different systems. This situation has certain implications for the availability and open use of data. The researcher's obligation is to share pseudonymised data rather than directly identifying information, but because these are often detailed data, they cannot be regarded as anonymised; see the Glossary for terms.

has been used in referring to the complex framework of issues, including legal and ethical ones, and even if those ethical and legal issues are resolved there are still problems associated with interoperability that may limit the use of data and these, too, must be resolved for data to be shareable.

There are also issues to resolve in sharing data when public and private sector interests collide, particularly when there may be different legal regimes and jurisdictions appertaining to intellectual property rights. Intellectual property rights, interoperability and other management issues for collaborative research are outside the scope of the present report but are being discussed extensively elsewhere, for example in the work of the ALLEA group on intellectual property rights[6] and the work of the International Science Council's Committee on Data[7].

Individuals may be concerned that data about themselves could be used to deny access to health care, employment, other social benefits or be used to inform other discriminatory purposes (such as permitted entry to a third country). High-level detailed data combined with socio-demographic variables can, in some instances, if combined with information that may be publicly available, render individuals directly identifiable with high probability. The public's trust in sharing health data with their doctors, hospitals and registries rests on the assumption that such data will be treated confidentially. Health data may be particularly sensitive because of the potential for use by those who do not share the individual's values and because the individual may not have access to, and control of, their own data. To reiterate, it is essential to address data privacy issues to maintain public trust in data-intensive health research (Kalkman et al. 2019a).

Data sharing is taken to include data transfer as well as remote access to data at its original location. This encompassing scope affects the degree to which technical tools can solve the data sharing issue although recent developments of privacy-enhancing technologies (PETs) begin to bring opportunities within reach to share personal data in a secure way (Royal Society 2019; European Commission 2020) and these opportunities will be discussed in section 4.5.

Earlier EU mechanisms governing the use of patient data had been criticised as overly complex, ambiguous and an obstacle to research. Reforms were introduced in the EU GDPR (https://gdpr-info.eu) to strengthen data protection safeguards, and to provide individuals with additional and stronger rights and control over their personal data. During the passage of this legislation, European academies and others worked hard to ensure that important exemptions were made to enable sharing of personal data for health and scientific research, while including proportionate safeguards to protect data subjects' interests (FEAM and EASAC co-signature documents 2014, 2015).

An introduction to the broader international context and the pivotal role of EU national ethics committees is provided in Box 2.

6| See https://allea.org/intellectual-property-rights/.

7| See https://codata.org for discussion of various issues, including intellectual property rights, data interoperability, usability, standards and repositories.

Box 2 | Protecting patients' rights by consent: information sources on global and EU developments

Consent may be sought for different purposes. A consent to participate in research is not necessarily the same as consent as a lawful basis for the data processing which is again different from a consent to data transfer to non-EU/EEA countries. These distinctions are critically important for the context of the present report and further work is required to determine the extent to which the different consent objectives overlap. Our scope does not extend to a discussion of different models of consent but information can be found elsewhere on principles (see, for example, Nuffield Council on Bioethics 2015) and on recent developments (e.g. dynamic consent; Budin-Ljosne et al. 2017)

Broad guidance, for example on privacy, and links to international and national legislation is provided by the World Health Organization (WHO)[8]. The European Patients' Forum has published[9] information for patients on rights to their data and how exceptions to consent for research purposes should be managed (with technical and organisational safeguards). Expert guidance for EU researchers, on ethics and data protection, including international data transfer, is disseminated by the European Commission[10]. The Council for International Organizations of Medical Sciences (CIOMS) in collaboration with WHO (CIOMS 2016) has also developed ethical guidelines for health-related research on consent (including consent for unspecified future use) for the collection, storage and use of biological material and related data. According to the GDPR, consent must be freely given, specific, informed and unambiguous[11].

For research, consent can and must be obtained in a fashion that enables participants to understand that the value of their participation will be maximised (Walport and Brest 2011). Thus, from this perspective, it can be regarded as unethical for an ethics committee to allow a study to proceed that does not maximise the potential value of results, by sharing where appropriate, while also protecting confidentiality. Information about national research ethics committees and relevant national and EU legislation is provided by the European Network of Research Ethics Committees[12].

Some of the issues for the "health data ecosystem", including consent, privacy and commercialisation, are still controversial and some commentators advise that the changing norms and frameworks in data governance and ethics requires rethinking of governance mechanisms (Sharon and Lucivero 2019).

Continues

8| See https://www.who.int/genomics/public/patientrights/en.

9| See footnote 5.

10| "Ethics and data protection", November 2018 at https://ec.europa.eu/info/sites/info/files/5._h2020_ethics_and_data_protection_0.pdf.

11| See https://gdpr-info.eu/issues/consent/.

12| See www.eurecnet.org/index.html.

Box 2 continued

Particular challenges affect large studies (see above), many of which run for many years. For these types of study, it is difficult to keep pace with new legal and technological developments potentially affecting consent. For instance, it might be difficult or impossible to renew the consent because participants might have died, it would be too expensive to contact and re-consent thousands of participants, and the response rate might be very low, resulting in possible biased samples. In view of these obstacles, many research studies opt for asking their ethical committee for permission rather than re-consenting participants.

Practical issues for consent within the GDPR for international transfer of data are discussed in section 3.4.1.

1.4 | GDPR limitations

The objective of a harmonising framework provided by the GDPR, for processing personal data for research purposes within the EEA, is welcome. However, significant challenges remain. A report from the European Parliamentary Research Service (EPRS 2019) examined issues for the impact of the GDPR on the research community; the EPRS analysis will be discussed in section 2.1. Another report from medical academies reviewed these challenges from the point of view of researchers and health-care providers (FEAM European Biomedical Policy Forum 2018).

In describing the history of the development of international data sharing norms, it has been observed that data protection law has become a significant hurdle to the sharing of personal data across jurisdictional borders (Phillips 2018). This cumbersome situation leads to risk-avoidance behaviour by privacy authorities, self-regulation by researchers and concerns about compliance by EU/EEA institutions. The consequences have been delay, wastage of resources and insufficient re-use of data: recreating data may be more practical than re-using data from others.

The transfer of personal data outside the EU/EEA is a particular problem (FEAM European Biomedical Policy Forum 2018), inadequately operationalised in the GDPR. In the present initiative, ALLEA, EASAC and FEAM focus on this concern, with particular regard to the lack of Article 46 transfer mechanisms that can be used for sharing personal data with public sector research institutions outside the EU/EEA. We acknowledge, of course, that there will still be other GDPR issues to resolve but our focus on international transfer mechanisms reflects a current and growing problem for EU/EEA researchers that has significant potential impact on patients.

1.5 | New opportunities to collect personal data

The need to solve this data transfer impediment is heightened by the maturity of many large European cohort studies and

databanks, where many volunteers and patients have provided both health data and biological samples and where it is now possible to harvest the knowledge from their contributions. Hundreds of thousands of individuals have provided data in the past – often decades ago – to many of these large ongoing epidemiological studies[13]. These projects have been possible thanks to substantial funding coming from European tax payers to build these resources and biobanks, and to the participation of European citizens, who have often provided their biological specimens. Even with these efforts, European data alone is not enough; to find important associations with disease, these data need to be combined with those from other regions.

Furthermore, there are growing opportunities to collect big datasets (the combination and analysis of very large and diverse sets of data). In addition to health data collected according to agreed protocols in multicentre research projects, there are increasing opportunities for examining data from electronic health records within national health systems, for example for population-based prospective studies and in health registries (Canova et al. 2019).

There are other emerging sources such as wearable monitoring devices and biobanks (Cleary 2020; G-Science Academies 2020; Shilo et al. 2020). Data from personal devices are normally collected on the basis of a personal agreement among members of a community, often as part of citizen science (Hecker et al. 2018). Data may also be collected from social networks exploring health-related aspects through record linkage and the application of algorithms (Capurro et al. 2014) but there is no formal control of these analyses when performed by private companies, often on databases stored in servers of unknown origin.

The demand for using these big datasets, if made accessible in a consistent format, is growing, for example in disease diagnosis and phenotyping, modelling and prediction of clinical outcomes, prioritising patients for early intervention strategies and assessing the influence of public health policies. A systematic literature review, commissioned by the Directorate-General for Health and Food Safety (DG Santé), to assess the added value created by using big data in public health (including examples from multi-national research) developed a range of recommendations for the EU, including priorities for awareness raising, data sharing, governance of access and use, and privacy regulation (Gesundheit Österreich Forschungs- und Planungs GmbH, 2016).

There is now considerable potential for generating and using "evidence about all aspects of health care to serve the needs of patients, clinicians and all other decision makers around the world" (Hripsak et al. 2015). Long-term structured anonymised health data are also quite important for further development of new areas of research such as artificial intelligence (FEAM European Biomedical Forum 2019).

13| See, for instance, the European Prospective Investigation into Cancer and Nutrition (EPIC) study, which involved 521,000 participants recruited across 10 European countries for almost 15 years: https://epic.iarc.fr/.

1.6 | Focusing on international personal data sharing for health research

As will be discussed in detail in the following chapters, data sharing between the EEA countries and outside remains difficult (Rabesandratana 2019; Ursin et al. 2019a). Less global sharing of data for research will result in significant disadvantages for research, innovation and health care and, thereby, hurt everyone, including Europeans: "If data and samples from Europeans are no longer part of the large international efforts, we will not learn whether what holds true in non-EEA collaborations also applies to European populations" (Ursin et al. 2019a).

Previous advisory groups to the European Commission have emphasised how health research depends on high-quality cross-border collaboration within Europe and beyond (EU Scientific Panel for Health 2016), but the impediments to wider international collaboration have remained unresolved. The EU is a world leader in health research and has global responsibilities. As mentioned earlier in this chapter in the context of COVID-19, it is imperative to improve data sharing in the public health community; however, as yet, the benefits from data-sharing relationships between high-income and low- and middle-income countries remain largely unrealised (Brack and Castillo 2015).

The specific issues discussed in our report are relevant to several of the Sustainable Development Goals (SDGs), notably SDG 3 (health) but also SDG 9 (innovation, with the target to enhance scientific research) and SDG 17 (partnership with targets for policy coherence, capacity building and access to science worldwide). More generally, the United Nations has discussed[14] how big data has the potential to support many of the SDGs if risks to individuals' rights are addressed. One of the risks, to privacy, requires data protection measures to be put in place. But another risk, inequality between "data haves" and "have nots", also needs tackling and this necessitates data sharing.

1.7 | Objectives of the ALLEA, EASAC and FEAM initiative

This project is the first tripartite collaboration between FEAM, EASAC and ALLEA. It benefits from the complementary expertise joined in these networks and the interdisciplinary perspectives created. Membership of the Working Group and project procedures are described in Appendix 1. FEAM and EASAC have a history of interest in optimising the use of health research data and in examining the issues surrounding the GDPR. ALLEA too has significant interests in the issues for sharing and using data (see, for example, ALLEA and Royal Society, 2019).

The project focuses on clarifying principles and options for GDPR reform, taking into consideration the legal, ethical and, in particular, privacy implications. The objectives are as follows.

- Articulate the value and impact of multi-national health research.
- Compare the potential of different solutions for ensuring sufficient transfer of health data outside the EEA.

14| "Big data for sustainable development" at https://www.un.org/en/sections/issues-depth/big-data-sustainable-development/index.html. See also "Cape Town Global Action Plan for Sustainable Development Data" at https://unstats.un.org/sdgr/hlg/Cape-Town-Global-Action-Plan/.

- Inform the European Institutions during the next evaluations of the GDPR[15] and the elaboration of guidelines especially with regard to the transfer of personal data outside the EEA so as to recommend improved procedures and standards for international transfer of high-quality data, safely and effectively, including by issuing specific guidelines for scientific research.

Our messages are directed to the following groups.

- Those who make or influence policy in the European Commission, European Parliament and Council of Ministers.
- Those who make or influence policy at the EU/EEA Member State level.
- Member academies and others in the scientific community, as well as other health stakeholders (including patient groups) who might benefit from optimal use and sharing of health data for research.
- Through our member academies, to the lay public, to Research Ethics Committees and to public health authorities.
- Those outside the EEA/EU who are interested in the international transfer of data for health research.

The following chapters present further information on the development of the GDPR, its current weakness and options for improving international mechanisms of data transfer. We set our analysis and recommendations into the broader context of other EU policy development for sharing health data and supporting open science. We also provide examples of where international sharing of health data for research has produced valuable inputs to policy, innovation and practice, which might not otherwise have been possible, and warn about what can now be lost without effective mechanisms of international transfer.

15| EU Commission, "Data protection rules as a pillar of citizens' empowerment and the EU's approach to the digital transition: two years of application of the General Data Protection Regulation", Communication from the Commission to the European Parliament and the Council, SWD(2020) 115 final, 24 June 2020. The GDPR mandates the EU Commission to periodically report on the evaluation and review of the GDPR. This first report was due after 2 years of the entry into force of the GDPR (2020), and the next reports will be issued every 4 years.

2 | Issues for scientific research raised by the GDPR

Individual examples of how sharing health data benefits patients, researchers and society are listed in Appendix 2. A systematic review of the outcomes of health-care research (Cruz Rivera et al. 2017) identified potential impacts in terms of patient benefits, informed policy-making, improved health services, and other societal and economic impacts.

2.1 | Development of the GDPR and issues for research

Some of the earlier activities of the academies of science and medicine, with others in the scientific, medical and patient communities, in helping to craft the GDPR and in assessing how likely it will meet the interests of those communities, were outlined in chapter 1. Original drafts of the EU Regulation set out a proportionate mechanism for protecting privacy, while enabling health and scientific research to continue. It included a requirement for specific and explicit consent for the use and storage of personal data but provided an exemption for research, dependent on certain safeguards, and recognised that research subjects' interests can be protected through strong ethical and governance mechanism, such as approval by a research ethics committee (Box 2) and good clinical practice frameworks. However, despite this exemption, the GDPR is making international research very difficult in practice, as will be described subsequently.

During the parliamentary passage of this draft regulation, the scientific, medical and patient communities expressed concern that the scope of the research exemption would be reduced. This would have put at risk significant European investments in genetics, cohort studies, biobanks and repositories, disease registries and the use of routinely collected data, and associated progress towards understanding society, health and disease (Joint Statement 2014[16]). Fortunately, these concerns were heard (Joint Statement 2014[17]); the European Parliament, European Commission and Council of Ministers demonstrated their shared commitment to research by finding a compromise position that enables vital health and scientific research to continue under the GDPR.

Nonetheless, research is at risk: there are continuing severe problems for the research community and some of these have been articulated previously by FEAM (see chapter 1 and Box 3). As well as specific issues to clarify for researchers in complying with the GDPR, more general issues are also raised for the relationship between researchers, citizens and data subjects in defining and optimising the interfaces between science and society (Starkbaum and Felt (2019) and

16| FEAM, Joint Statement, 2014. https://www.feam.eu/wp-content/uploads/HealthcareCoalitionOnDataProtection_2014_jointstatementPUBLISHED-2.pdf.

17| FEAM, Joint Statement, 2014. https://www.feam.eu/wp-content/uploads/LIBEreportJointStatementFebruary2014-1.pdf.

Box 3). However, broader issues for involving patients in the design and conduct of health research are beyond the scope of the present report.

There is also concern at the multiplicity of laws (e.g. GDPR, clinical trials regulation, medical treatment legislation) and journal editorial practices dealing with overlapping issues and potentially generating conflicting requirements for researchers.

Box 3 | General policy priorities: critical remarks and recommendations adapted from the FEAM European Biomedical Policy Forum (2018)

- At the level of implementation in Member States, there is still a need to produce a fully harmonised framework in some areas of research, subject to further information on how the GDPR has been implemented in the scientific research field in each Member State.
- Uncertainty and ambiguity in interpreting GDPR provisions and obligations might lead to sub-optimal use of health data for research, including potentially abandoning or not initiating projects, or potentially reducing their scope owing to the fear of non-compliance. Uncertainties may also lead to higher costs for research institutions in terms of requiring additional support for compliance activities.
- At the European Commission level, there would be value in constituting a cross-Directorate-General multi-stakeholder group to monitor the implementation of the GDPR in research with health data, and as a mechanism to receive feedback from the medical and science communities.
- More broadly, there must be continuing analysis on whether GDPR strikes the right balance between privacy and research interests (see further discussion by Bentzen and Hostmoelingen (2019) and Mascalzoni et al. (2019)).

EPRS assessment of current concerns by the research community (EPRS 2019) also highlights the need to balance the rights to academic freedom (to engage in innovative research), which includes the privilege to analyse personal data (Ursin et al. 2019b), and to protect against misuse of sensitive personal data. The comprehensive EPRS analysis notes that one of the particular concerns expressed by the research community is the difficulty of cross-border transfer of data to non-EU countries. Although the EPRS 2019 assessment itself concluded that the potential impact of the GDPR on data transfer outside the EU/EEA is neither positive nor negative, the evidence that we will cite in the present report testifies to a growing problem, and our recommendations call for reform. Solving this issue of safe data transfer requires new commitment to global thinking about the opportunities and challenges, to encourage the European Commission and other countries to agree on conditions that will protect the privacy of individuals, while acknowledging that some parties are not subject to European law.

2.2 | What international health research is at risk?

The consequences of not solving this problem are already impeding and will continue to impede health research in many critically

important fields. The Working Group discussed key objectives for sharing data with researchers outside the EU/EEA that included accumulating data on adequate patient numbers to understand how to prevent or treat rare diseases and specific subtypes of common disease; identifying complex disease patterns including combining data from laboratory results and other patient data; evaluating the determinants of therapeutic success and failure, and side-effects; and comparing local data sets to inform local health services and adjust their policies. Multi-national research collaborations and large consortia are increasingly necessary and anonymisation of shared data is rarely possible, for example in large-scale epidemiology and registry data with a high number of variables from different countries. The current problem affects both sending data and granting access to data repositories. The impediment is not confined to research objectives: for example, public health monitoring of cancer survival, including stage-specific rates, requires detailed data, which currently cannot be shared outside the EEA. Some of these fields, identified as EU priorities in the early discussions to scope this academies' project, are listed in Box 4.

Box 4 | Examples of health research priorities for sharing data

- Rare disorders, including rare cancers and subtypes of common cancer. For example, the US National Cancer Institute Cohort consortium[18], a consortium of more than 60 cohorts of 50,000–100,000 individuals across the world, has provided data for a vast amount of cancer research. However, much of the EU/EEA contribution to this collaboration has come to a standstill given the inability of European researchers to share data with the US National Institutes of Health (NIH) post-GDPR.
- More generally for cancer registries and large international cancer epidemiology collaborations, where there are no obvious transfer mechanisms for sharing EU personal data with the WHO International Agency for Research on Cancer[19].
- Psychiatric disorders and autism. For example, the Psychiatric Genomics Consortium has experienced problems in sharing EU–US data for rare subtypes in psychosis, bipolar and eating disorders.
- Infectious disease as highlighted by the COVID-19 pandemic (where only initial transfers were covered by the derogation; see section 3.4). Also, antimicrobial resistance, where data on patient flow within and between countries is needed, together with sharing for re-use of national data between countries for modelling.
- Microbiome research, where there are privacy issues because each individual's microbiome is unique.
- Genetic research, also with privacy issues. Large studies are essential to have sufficient numbers within subgroups of disease and combinations of genetic variants. The problem is illustrated by the recent example of the Pan-Cancer Analysis of Whole Genomes, which could not establish a truly international cloud-service because of EU restrictions on data transfer across borders (Phillips et al. 2020).
- Disease prevention strategies, including vaccination and child health programmes. For example, research that requires sharing data from population-wide screening programmes of maternal health and child development, including newborn screening for congenital diseases.

18| https://epi.grants.cancer.gov/cohort-consortium.

19| See Global Cancer Observatory, https://gco.iarc.fr.

Other examples of problems in sharing data have been documented in the literature: see, for example, EPRS (2019) for a review, Rabesandratana (2019) for an example of problems experienced between the National Institute for Health and Welfare, Finland and the US NIH, and Mitchell et al. (2020) for a discussion of the issues for genomic data[20].

A publication from the USA (Peloquin et al. 2020), with particular regard to biobanks, also found disruptive and avoidable challenges introduced by the GDPR, including problems in cross-border transfer. A perspective from the US NIH emphasises the unintended consequences of the GDPR in hampering international researchers, for example in NIH collaborative studies in cancer and diabetes (Eiss 2020). Recent discussion between US and EU/EEA public institutions may help to clarify uncertainties, develop examples of good practice and set precedents for formal data use agreements under the GDPR; resolution of the problems is as urgent for the global scientific community as it is for the European scientific community. The problem may become one of considerable magnitude: it was estimated in 2019 that there were approximately 5,000 collaborative ongoing projects between the US NIH and EEA countries[21], most started before the GDPR, but there is no mechanism within the GDPR to update these previous agreements (see subsequently for more details). The proportion of these projects involving data sharing is not known but it is probably high.

There is a further problem in that the start of some already-funded research studies involving third-country collaborations has been delayed because of GDPR issues.

20| Further case studies are included in Appendix 2.

21|http://www.iscintelligence.com/archivos_subidos/robert_eiss_gdpr_us-eu_cooperation_in_biomedical_science_isc_gdpr_seminar_19_nov_2019.pdf.

3 | Solutions provided by the GDPR

The GDPR provides for a layered approach[22] to tackling the current diversity of data protection procedures around the world[23]. To resolve the current GDPR problem in the short-term is not a matter of adding extra mechanisms for international transfer, but rather to ensure that the current tools can be made to work well. As a general point, while the GDPR provisions recognise different scenarios, their emphasis has been on private sector operators, not the public sector.

3.1 | Adequacy (GDPR Article 45)

Free movement of data from the EEA is allowed if there is an "adequacy" decision for the recipient. The requirements for a country to meet adequacy standards are strict (EPRS 2019) and depend on whether strong privacy rules are already applied within that country. So far, the European Commission has recognised only a few countries as having adequate protection (i.e. Andorra, Argentina, Canada (only as it concerns commercial organisations), Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland and Uruguay). Therefore, an adequacy decision for major research-intensive countries such as China, Australia, the USA and South Africa is currently lacking and very unlikely to happen for countries lacking a legal framework for privacy protection such as Australia or China.[24]

For the USA, no adequacy decision has so far addressed the needs of public sector researchers. Moreover, while initially covering the private sector, both the EU Commission Safe Harbour decision issued in 2000[25] and its replacement, the EU Commission adequacy decision on the EU-US Privacy Shield[26], have now been overturned by the European Court of Justice.

In two landmark EU data protection law cases brought by Austrian lawyer and data rights activist Max Schrems, the European Court of Justice (ECJ) first invalidated the EU–US safe harbour agreement ("Maximiliam Schrems v. Data Protection Commissioner" or Schrems I)[27], and secondly invalidated Decision 2016/1250 on the adequacy of the protection provided by the EU-US Data Protection Shield ("Data Protection Commissioner v. Facebook Ireland and Maximilian Schrems"

22| European Data Protection Supervisor guidance on international transfers is at https://edps.europa.eu/data-protection/data-protection/reference-library/international-transfers_en.

23| See mapping by Commission Nationale de l'Informatique et des Libertés, November 2019, at https://www.cnil.fr/en/data-protection-around-the-world.

24| https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en.

25| European Court of Justice 2000/520/EC: Commission Decision of 26 July 2000 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the safe harbour privacy principles and related frequently asked questions issued by the US Department of Commerce (notified under document number C(2000) 2441).

26| Commission Implementing Decision (EU) 2016/1250 of 12 July 2016 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the EU-U.S. Privacy Shield, OJ 2016 L 207.

27| ECJ, Case C-362/14, Schrems v. Data Protection Commissioner, ECLI:EU:C:2015:6506, Judgement of 6 October 2015.

or Schrems II)[28]. In this later case, the ECJ also ruled that the EU Commission's decision on standard contractual clauses (SCCs; see below) is valid. Nonetheless, the ECJ reminded that the validity of such decision depends on the inclusion of effective mechanisms, including supplementary measures, to ensure compliance with the protection required by EU law. This means that transfers of personal data outside the EEA are to be suspended or prohibited in the event of a breach of such clauses or the impossibility to comply with them.

More generally, these cases have also put in evidence the viewpoint of several non-governmental organisations, which are critical of what they regard as an apparent insufficient protection of privacy under the current data protection framework.[29]

3.2 | Appropriate safeguards (Article 46)

Article 46 of the GDPR establishes that "in the absence of a decision pursuant to Article 45(3), a controller or processor may transfer personal data to a third country or an international organisation only if the controller or processor has provided appropriate safeguards, and on condition that enforceable data subject rights and effective legal remedies for data subjects are available". Among the potential appropriate safeguards listed in Article 46 are SCCs, administrative arrangements between public authorities or bodies, or specific contractual clauses. Administrative arrangements and specific contractual clauses (or bespoke contracts where at least one of the contractual parties is private) require the authorisation of the competent supervisory authority.

3.2.1 | Standard contractual clauses (SCCs)

The most common instrument for appropriate safeguards is the standard contractual clause (SCC) (Article 46(2)(c) of the GDPR) but such clauses cannot be amended in any manner. As such, they have been found to be inflexible, and have not been agreed by major potential research partners, such as the United Nations and the US NIH, because of conflict with US federal laws (Peloquin et al. 2020). While SCCs have worked well for private universities and private institutions, obstacles to collaborations with public research institutions remain. These include the rules governing judicial redress and indemnification, and non-EEA countries' archiving laws, which pose additional obstacles[30]. In November 2020, the EU Commission published a draft implementing decision on SCCs for the transfer of personal data to third countries pursuant to the EU GDPR, along with new SCCs, which are expected to be adopted in early 2021. However, the new SCCs left many of the issues for research unsolved.

28| ECJ, Case C-311/18, Data Protection Commissioner v Facebook Ireland Limited and Maximillian Schrems, Judgement of 16 July 2020.

29| See among these, noyb (My Privacy is None of Your Business) founded by Max Schrems who initiated both of the above-mentioned cases, and "La Quadrature du Net", who initiated earlier complaints against Google, Apple, Facebook, Amazon and Microsoft, which were later made available as templates for other citizens to file complaints across the European Union, and which case was recently decided by the ECJ (joined Cases C-511/18, La Quadrature du Net and Others, C-512/18, French Data Network and Others, and C-520/18, Ordre des barreaux francophones et germanophone and Others (often referred as "La Quadrature du Net and Others"), Judgement of 6 October 2020.

30| See comments submitted by the Norwegian Institute of Public Health and the Cancer Registry of Norway to the EDPB in response to the public consultation on Guidelines 2/2020 on Articles 46(2)(a) and 46(3)(b) of Regulation 2016/679 for transfers of personal data between EEA and non-EEA public authorities and bodies: https://edpb.europa.eu/sites/edpb/files/webform/public_consultation_reply/edpb_guidelines_niph_crn_comments_20200518.pdf.

It is important to highlight that changes to the standard wording of SCCs that enable all prospective partners to agree would provide a solution to the current problem of international transfer of data. This opportunity was missed in the context of recently proposed revision of the SCCs, which clarified a few issues about archiving laws, but left out the issue of redress and indemnification.

3.2.2 | Administrative arrangements between public authorities and the EDPB

Appropriate safeguards may also be provided by means of a legally binding and enforceable agreements between public bodies (Article 46(2)(a) of the GDPR), or (subject to authorisation from the competent supervisory authority) through provisions inserted into non-legally binding administrative arrangements between public bodies which include enforceable and effective data subject rights (Article 46(3)(b)). The EDPB recently consulted on guidelines 2/20 appertaining to these administrative arrangements[31]. The Norwegian Institute of Public Health and the Cancer Registry of Norway submitted comments drawing attention to issues for redress mechanisms (e.g. liability for data breaches, ability to sue), which is an impediment to data transfer to the USA; and archiving laws, where it is unclear if third countries' archiving requirements can be fulfilled under GDPR Articles[32]. For these reasons, this type of arrangement is not a viable solution to solve the challenges of international transfers, especially with regard to US federal institutions.

3.2.3 | Bespoke contracts

Appropriate safeguards may also be provided through "the use of contractual clauses between the controller or processor and the controller, processor or the recipient of the personal data in the third country or international organization" (Article 46(3)(a)). Such bespoke contracts offer another avenue for public research organisations to transfer health data outside the EEA; nonetheless, they need to be subject to the authorisation of the competent supervisory authority.

While bespoke contracts potentially offer a valid solution for international transfers, their use has been limited by the lack of guidance from the EDPB on the specific requirements, which has led data protection authorities to refrain from establishing a process for reviewing such clauses[33] (Peloquin et al. 2020).

3.3 | Codes of conduct

Some parts of the GDPR have now been interpreted in terms of codes of conduct but the EDPB has yet to issue guidelines relating to data transfer mechanisms. Judging by the codes of conduct specified for within-EU transfer of data, the guidelines are likely to be complex and require creation of an independent

31| See EDPB proposal and public responses at https://edpb.europa.eu/our-work-tools/public-consultations-art-704/2020/guidelines-22020-articles-46-2-and-46-3-b_en.

32| See comments submitted to the EDPB at https://edpb.europa.eu/sites/edpb/files/webform/public_consultation_reply/edpb_guidelines_niph_crn_comments_20200518.pdf.

33| See, for instance, the position of the UK Information Commissioners' Office: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/international-transfers/.

monitoring board[34] which may delay research and add costs for the researcher. There is an important opportunity here for the science community to communicate how codes of conduct can be made particularly relevant and supportive for research needs. While such mechanism would allow the scientific community to build a bottom-up solution, it is important to consider that their initiation validation, approval and implementation will require considerable time and resources, therefore creating impediments to scientific research.

3.4 | Derogations (exceptions, Article 49)

Article 49 of the GDPR establishes the conditions upon which personal data may be transferred to a third country or an international organisation in the absence of an adequacy decision or of appropriate safeguards. Nonetheless, as required by Article 44 of the GDPR, protections need to be applied in such a way as to ensure that the level of protection of natural persons guaranteed by the GDPR is not undermined (for instance, the use of derogations may never lead to a breach of fundamental rights).

Under the "layered approach" envisaged by the EDPB as well as by its predecessor (Article 29 Working Party), derogations may only be used in the absence of an adequacy decision or availability of appropriate safeguards. Moreover, derogations included in Article 49 are exemptions, and as such according to inherent European law principles, they must be interpreted restrictively to avoid that the exception becomes the rule.[35]

3.4.1 | Explicit consent

Explicit consent means the data subject has consented to the proposed transfer, after having been informed of the possible risks of such transfers for the data subject owing to the absence of an adequacy decision and appropriate safeguards (i.e. tantamount to relinquishing protection). Even if this form of consent were to be appropriate for future research (possibly only for small-scale studies; EPRS 2019), previous research consents are unlikely to have included this stipulation. And, at the outset, researchers may not know the countries to which they would subsequently wish to transfer data.

3.4.2 | Important public interest

Pursuant to Article 49 of the GDPR, there are specific situations under which a transfer can be exceptionally allowed. However, this must already be specified in the law of the Union or a Member State and the threshold is likely to be high, for example exchanging data in responding to a pandemic, and the Danish Data Protection Authority advised the example of Ebola. In the recently published guidelines on the processing of data for the COVID-19 outbreak, the EDPB has reiterated that such exemptions, both "transfer necessary for important reasons of public interest" (Article 49(1)(d)) and "explicit consent" (Article

34| See https://edpb.europa.eu/our-work-tools/our-documents/nasoki/guidelines-12019-codes-conduct-and-monitoring-bodies-under_en.

35|https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_2_2018_derogations_en.pdf.

49(1)(a)) "must be interpreted restrictively and on a case-by-case basis".[36]

Furthermore, although Article 49 can be used for transfer of data relating to COVID-19, this applies to initial transfer only, not repetitive transfer of data (that would fall under Article 46). For consistency, it would be helpful for further guidance to be provided by the EDPB to the national data inspectorates on how public interest exceptions should be decided. The principle of restrictive interpretation is reasonable because if too many exceptions were allowed then other transfer mechanisms might be ignored. However, if it is too restrictive, important benefits might be lost.

3.5 | Supplementary measures

An additional layer of complexity is added by the requirement to use supplementary measures to ensure that the standard of protection for the data is essentially equivalent to that provided by EU law, when such transfers rely on the safeguards established by Article 46 of the GDPR. In such cases, data exporters must identify and adopt supplementary measures.

The use of such supplementary measures for international transfers as recently highlighted by the Scherms II decision entails additional difficulties for scientific researchers. Although the EDPB has recently issued guidance on this topic[37], this remains very high-level, and as many questions remain unanswered, national data protection authorities and local data controllers could be prone to precautionary limiting international transfers. Among the problems raised for scientific research is that the EDPB has set up a level of pseudonymisation that exceeds the requirements of the GDPR, therefore creating additional hurdles that might be impossible to meet for genetic and high-dimensional data (e.g. genetics research), or for other data containing a large number of public linkage sources[38].

3.6 | Key conclusions from this chapter

Thus, in summary, according to the provisions in the GDPR, the data subject can best be protected through an adequacy decision (Article 45) or through appropriate safeguards (Article 46). Consent is a weaker protection for the individual (see also Box 2) and should only be used as a transfer mechanism in the absence of an adequacy decision and safeguards (Article 49). Hence, a transfer mechanism that provides an appropriate safeguard will best serve the data subject—and it also happens to be a more practical solution for ensuring data flow for research. However, as described above, there is a lack of non-consent-based transfer mechanisms that can be used for sharing personal data with public institutions in the USA and elsewhere and providing access to data repositories[39]. It is noteworthy that the previous (i.e. pre-Schrems II) mechanism,

36| See https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_en.pdf.

37|https://edpb.europa.eu/sites/edpb/files/consultation/edpb_recommendations_202001_supplementarymeasurestransferstools_en.pdf.

38|https://www.nshg-pm.org/NSHG-PM-takes-up-post-schrems-II-health-data-sharing. For the full text of the response to the public consultation submitted in December 2020, see: https://edpb.europa.eu/sites/edpb/files/webform/public_consultation_reply/comments_from_niph_and_uio_on_recommendations_01-2020_on_measures_that_supplement_transfer_tools.pdf.

39| See also further detailed discussion by ISC "the application of GDPR to biomedical research: stakeholder advisory opinions to assist regulators": http://iscintelligence.com/archivos_subidos/input_paper_on_gdpr_challenges_for_research-77623379-v15.pdf.

offering the ability of private companies to use Privacy Shield and SCCs whereas public sector researchers could not, gave a large advantage to the private sector, capitalising on data produced in the public sector with public investment.

Of course, apart from the lack of clarity and support mechanisms in the GDPR, there are additional impediments to data exchange for research—including those deriving from lack of research capacity and resources in recipient countries. Nonetheless, in African case studies on biobank data transfer, the GDPR was identified as an obstacle to research and to research capacity building in low- and middle-income countries (Slokenberga et al. 2019). It is important to emphasise that while mechanisms such as SCCs work better for private institutions, they do not work for many public or governmental institutions, such as federal institutions in the USA, which include major research partners (or funders) of many European researchers. With more than 5,000 ongoing collaborations involving EU researchers and the US National Cancer Institute, this represents a critical challenge for the research community[40]. It is very likely that many of these collaborations involve actual transfers, and therefore the absence of a workable Article 46 transfer mechanism to federal institutions in the US remains a very important hurdle for effective collaborations.

The recent EDPB guidance on supplementary measures adds useful discussion, but it also left out many problems and uncertainties. As an example, assessing non-EEA laws is a daunting task for individual research institutions, and this requirement may represent a serious hindrance to data sharing moving forward. Moreover, the problems related to the use of SCCs remain with the recently revised version of November 2020.

40|http://www.iscintelligence.com/archivos_subidos/robert_eiss_gdpr_us-eu_cooperation_in_biomedical_science_isc_gdpr_seminar_19_nov_2019.pdf.

4 | Issues at stake for international data transfer

Different countries worldwide have different laws, affecting the operation of the GDPR. The ALLEA, EASAC and FEAM initiative welcomes the recognition by the European Commission that there are continuing issues relating to the international transfer of personal data to third countries[41] but we urge the European Commission to listen to concerns from those – such as the public sector research community – who are not part of their authorised GDPR multi-stakeholder group. In our view, the discussion must also extend beyond the Directorate-General for Justice and Consumers. If the potential benefits of data-driven research and care are to be realised, there needs to be better interaction with multiple Directorates-General, including the Directorate-General for Research and Innovation (DG Research) and the Directorate-General for Health and Food Safety (DG Santé), to communicate the biomedical community's concerns and priorities (FEAM European Biomedical Policy Forum 2018).

4.1 | How might the GDPR be improved in the short-term to support reliable transfer mechanisms?

In terms of the transfer mechanisms discussed in chapter 3, in the opinion of ALLEA, EASAC and FEAM the following are recommended.

- A core problem is that the GDPR provisions have not resulted in operational data transfer mechanisms (outside private institutions). The best option would be to find a workable solution under Article 46 and revise the wording of SCCs (or create an additional specific SCC for scientific research purposes) so that they can be agreed by other research partners. The urgent challenge is to find a simple solution that is safe and does not conflict with other countries' laws. This lack of workable legal mechanisms is currently affecting a large number of ongoing collaborations, including more than 5,000 collaborative projects with the US NIH (see above)[42]. We suggest that the EDPB could consider providing guidelines on when Article 49 derogations can be used for existing transfers.
- It is also urgent for the EDPB to produce guidelines (accompanied by examples for health research) on codes of conduct and certification and for bespoke contracts where one or both partners is not a public body.
- The EDPB should also urgently revise current draft guidelines (with concrete examples) for administrative arrangements and bespoke contracts between public bodies so that they do not conflict with national laws in non-EU/EEA countries.

41| European Commission (2020) and European Commission Roadmap "Report on the application of the General Data Protection Regulation", Ares (2020)1873825-01/04/20.

42| http://www.iscintelligence.com/archivos_subidos/robert_eiss_gdpr_us-eu_cooperation_in_biomedical_science_isc_gdpr_seminar_19_nov_2019.pdf.

- Other options are less clear. Further consideration should be given to the issues for the EDPB in developing guidelines on other international transfer mechanisms, including when personal data can be transferred because of the "important public interest" condition. The problem with derogations is that the responsibility is placed on the research participant rather than the recipient of the data, but further discussion on the public interest derogation (and whether it can be expanded) would be useful.
- In the wider international context, the EU should lead discussions to encourage all countries to adopt consistent privacy rules that safeguard the privacy of subjects and support collaboration in a safe manner (recognising that currently there is substantial variation in guidelines on ethical principles and norms; Kalkman et al. 2019b), whereby adequacy decisions can become broader in scope. Global convergence in privacy rules, together with creation of a "Data Protection Academy" to disseminate best practice, would bring new opportunities to protect European citizens while facilitating international data sharing (European Commission 2020) and encouraging reciprocity with other countries for sharing their data.

4.2 | UK status post-Brexit

An additional concern for the international science community arises in consequence of the withdrawal of the UK from the EU. At the time of writing this report (March 2021), the UK becomes a "third country" outside the EEA/EU. While current rules have been kept for 6 months (until 30 June 2021), future exchanges of data will depend on whether an adequacy decision from the European Commission could be made within this timeframe[43].

On 19 February 2021, the EU Commission started the process of adopting an adequacy decision for transfers of personal data from the EU to the UK[44] (data flows from the UK to the EU have been ensured through UK legislation applying since 1 January 2021). A final adequacy decision would still require a positive opinion from the EDPB and from a committee of representatives of EU Member States. In announcing its decision, the EU Commission highlighted how EU law has shaped the UK's framework for data protection during the past decades. A final adequacy decision would be valid for a first period of four years, and could be renewed if the UK protection for personal data continues to be deemed adequate.

Presumably the UK could either try to ensure that its laws continue to meet EU requirements and maximise the chance of an adequacy decision by the Commission to be maintained over time, or it could decide to base its data protection standards on other jurisdictions (if these are deemed more important for research objectives) (Taylor et al. 2018). This major issue for UK research, which also affects personal health data sharing in EU-funded projects with the UK as a partner, requires further consideration.

43| https://www.nature.com/articles/d41586-021-00009-y.

44|https://ec.europa.eu/commission/presscorner/detail/en/ip_21_661.

4.3 | Alternative models for the EU to consider in the longer term

The WHO is aiming to become a hub for how to handle health data (Shaffer 2020), including through the work of its Digital Health Technical Advisory Group. While it may be possible for the EU to align itself with future WHO-established health data procedures reflecting general principles (such as privacy), the current framework for international transfers within the GDPR has interfered with data sharing between the EU and international organisations, including the WHO (see Appendix 2 with examples from the WHO International Agency for Research on Cancer).

There may also be opportunities to develop an international code of conduct, as proposed for example in genomics, to clarify how researchers can comply with various laws, for example GDPR and the US Health Insurance Portability and Accountability Act (Phillips et al. 2020). The Biobanking and Biomolecular Resources Research Infrastructure-European Research Infrastructure (BBMRI-ERIC) initiative to develop a code of conduct for health research[45] is a useful step. Others have argued for unrestricted use of public genomic data, for example Amann et al. (2019), but this is seen as contradicting current efforts in data governance and raises problems for the nature of consent as well as researchers' and research funders' expectations (Nicol et al. 2019). In the longer term, if adequate data protection could be guaranteed in all territories, then reduced regulation of international data sharing could be contemplated (Phillips 2018), as intimated in section 4.1.

45| http://code-of-conduct-for-health-research.eu.

4.4 | Other EU developments

Recent developments in forming the European Open Science Cloud[46] bring new opportunities to build the infrastructure, standards and services to handle sensitive clinical data. It is important for all relevant EU developments to be aligned. The implications of constituting the EU Health Data Space also need to be considered. Development of the European Health Data Space recognises the importance of facilitating the exchange and sharing of health data across Europe, both the primary use of data for health-care delivery and the secondary use of data for research and policy-making. It would seem opportune to consider how these objectives can be supported by improved procedures for the safe international sharing of health data with researchers outside the EEA/EU.

In fact, the recently announced EU Data Strategy, which calls for the creation of EU Data Spaces, including a Health Data Space, pointed out how "sensitive data (e.g. health data) in public databases is often not made available for research purposes, in the absence of capacity or mechanisms that allow specific research actions to be taken in a manner compliant with personal data protection rules"[47]. Among the forthcoming actions foreseen in the EU Data Strategy is the development of a legislative framework for the governance of common European data spaces that address this and other related problems.

46|https://ec.europa.eu/digital-single-market/en/european-open-science-cloud.

47| EU Commission, A European strategy for data, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2020) 66 final, 19 February 2020.

4.5 | Technology options: privacy-enhancing technologies (PETs)

Privacy-enhancing technologies (PETs) are a wide range of technologies aimed at mitigating security and privacy risks. Some of these technologies enable the use of data without giving access to all or part of the data to other people, therefore potentially offering solutions to either enhancing privacy during data sharing, or bypassing the need for transferring or sharing data[48]. However, there is great variation among available PETs: most involve the transfer of personal data and may be considered as supplementary measures to enhance privacy. Only those PETs that do not involve the transfer of data may be considered an alternative solution for the international transfer of data, keeping in mind that provision of remote access also constitutes data transfer.

Academies have previously looked at data science and technologies, including new secure processing approaches. In 2017, the Royal Flemish Academy of Belgium for Science and the Arts published a report on Data Science and Healthcare, which also covers PETs (Verdonck et al. 2018). A report by the Royal Society (2019) provides further detail on PET capabilities and limitations and the potential opportunities for privacy-preserving data analysis.

The implementation of the GDPR provides impetus for further development and uptake of PETs. The GDPR's approach of "data protection by design and by default" (Article 25) encourages the use of technical (and organisational) measures to ensure the implementation of data protection principles such as data minimisation (i.e. ensuring that only a minimum required amount of relevant data for a given application is collected or stored). In this sense, PETs are increasingly viewed as a potential set of safeguards to be used as an alternative to, or complementary to, other approaches such as encryption or pseudonymisation (Royal Society 2019).

When developed appropriately and at the right level of technical maturity, some PETs could offer advantages, such as the possibility to reduce the privacy risk associated with data processing mechanisms, and could open up possibilities for the use, access, analysis and sharing of data that would otherwise be impeded because of privacy concerns, therefore potentially enabling further data uses. Being at the forefront of PET development will be an asset that makes Europe an attractive location for international collaboration.

Some PETs combine de-identification techniques that do not necessarily render the data anonymous (and therefore outside the scope of the GDPR), but may in certain cases provide excellent supplementary measures (Kaissis et al. 2020).

While no PET seems to be currently available to offer all technical solutions at the same time, available PETs have been used for different purposes with the following reported advantages:

- providing secure access to private datasets;
- enabling joint analysis on private data held by several organisations;

48 | See Rina Shainski and William Dixon, "How privacy enhancing technologies can help COVID-19 tracing efforts", World Economic Forum, 22 May 2020, at https://www.weforum.org/agenda/2020/05/how-privacy-enhancing-technologies-can-help-covid-19-tracing-efforts/ and World Economic Forum, The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value, September 2019, http://www3.weforum.org/docs/WEF_Next_Gen_Data_Sharing_Financial_Services.pdf.

- outsourcing computations on private data to the cloud in a secure manner;
- de-centralising services that rely on user data.

Among the PETs offering different degrees of solutions to address these issues with potential implications for health research are the following: (1) homomorphic encryption allowing data to be encrypted before it is shared; (2) differential privacy adding noise to an analytical system to render it impossible to trace back individual inputs; (3) federated analysis allowing parties to share the insights of their analysis without sharing the data; (4) confidential computing using hardware-based techniques to isolate data, specific functions or the entire applications of an operating system, virtual machines or other processes; (5) secure multi-party computation spreading data across multiple parties so that no individual party is able to see the complete set of inputs. While not a PET, blockchains provide an open, distributed ledger that can record transactions between several parties and can be combined with PETs to be made more privacy-preserving. More details, as well as a full description of these techniques with some of their advantages and limitations, are included in Appendix 3.

Some limitations of PETs include the risk of losing some utility and accuracy in the data, the ability to scrutinise the data before using it; or the use of significant computation resources as well as the financial costs of implementing some of these PETs. Finally, while some of these technologies are ready to be used, many of them have not reached technological maturity and are still at an early stage of development. The above-mentioned technologies have been used to handle health data in several projects that are described in Appendix 3.

Remote access, albeit still considered a transfer, poses additional problems. For instance, cloud computing provides the means for local on-demand access to resources but a centralised international cloud networks raises issues over data residency and governance, as well as the more general question of whether the system can be used in a trustworthy manner. These issues are attracting political attention[49].

Overall, the development and adoption of PETs seem to be a work in progress. Promising areas include decentralised data storage and federated learning systems, which could potentially change current standards for data sharing and for centralised storage of data. They have also been used in the biomedical area including medical imaging and genomics. Some of the limitations of current PETs could be counterbalanced by the use and further development of other techniques, including encryption (Kaissis et al. 2020). Nonetheless, while PETs might become more useful in the future, for the moment they serve only limited purposes, in particular as supplementary measures.

49| See the project by the UK Parliamentary Office of Science and Technology examining governance issues for security, resilience and data ownership, at https://post.parliament.uk/work-programme/cloud-and-edge-computing. Also, the UK NHS (April 2018, pre-Brexit) advised that hosting patient health data in public cloud services is only possible when hosted by the EEA or a country with an adequacy decision: https://digital.nhs.uk.

5 | Conclusions and recommendations

1 | Health research is crucial for all

It is in the interests of EU citizens and society as a whole for data to be shared for research. Health research is very important for patient benefit, population health, development of health-care systems, and for social cohesion and stability. It is an important ethical obligation to promote the health of citizens through appropriate research. Data from EU citizens must be shared for international research, to ascertain if new findings elsewhere also apply to EU populations.

2 | Sharing pseudonymised personal health data for public sector research is important

Sharing pseudonymised personal health data for research makes best use of limited resources and must be encouraged to maximise the individual and societal benefits to be obtained from the contribution of patients and volunteers to research. This is an important contributor to sustaining well-founded public trust and confidence, grounded in a broadly agreed social contract.

3 | Sharing of data safely and efficiently, as part of responsible science, must take account of privacy concerns

It is vital that health data are shared safely and effectively and that patient/volunteer privacy is respected. Researchers are accountable for their research and this includes their actions in sharing data. Building public trust in data sharing depends on building public trust more generally in the conduct and oversight of science and it is important to take account of patients' views. Academies have previously made a wide range of recommendations to develop and maintain responsible science and promote public awareness of those responsibilities for the science community[50].

4 | Implementation of the GDPR has created impediments to share health data for research internationally

The objective of a harmonising framework provided by the GDPR, for processing personal data for research purposes within the EEA, is welcome. However, there are significant hurdles for sharing data with researchers outside the EU/EEA, including EU collaborative research studies, when other countries may not have equivalent legal frameworks for data protection. The measures introduced by the GDPR to transfer data outside the EU/EEA do not provide workable mechanisms for sharing personal data with public sector research institutions outside the EU/EEA. It is essential to introduce an operational data transfer mechanism, functioning without further delay.

50| See the work by the InterAcademy Partnership on responsible science and research integrity: https://www.interacademies.net/news/world-science-academies-release-report-promote-research-integrity.

5 | Finding a solution to overcome the barriers in sharing data is urgent

Our preferred solution is to find a workable – adequate and safe – solution under Article 46 of the GDPR, with operational guidance provided as a matter of urgency by the EDPB. Moreover, given the diversity of data transfers governed by the GDPR, it would be very helpful if guidelines were accompanied by tangible examples from the health sector for good practice, including guidelines on how existing transfers and ongoing collaborative research can continue. For appropriate safeguards, a solution must be identified that is not in conflict with US or other laws outside the EU/EEA and that provides a redress mechanism for EU/EEA countries. If such a solution is not achieved then there is risk of inadvertent consequences whereby researchers are tempted to circumvent the GDPR by calling data anonymised, when they are not, or by trying for a derogation under Article 49, when that gives less protection to the subject providing personal data.

6 | Solutions should also include increased commitments to enable the use of shareable data

Even when appropriate mechanisms for transferring data are available, other methodological and technical quality issues need to be resolved to enable interoperability in the use of data. These challenges require greater attention across the research community.

7 | PETs are relevant but do not provide the solution to the difficulty of operationalising the GDPR

There are rapid developments in PETs that offer potential to improve data protection and data sharing agreements. It is important to assess and implement good practice in the use of such technologies now, even though they may be immature, as well as to look to the future for further developments that may overcome some of the current limitations. Continuing advances in technology development can be expected to increase data security and mitigate risk but they would not circumvent the requirements of the GDPR.

8 | Recommendation for continuing monitoring and assessment

EU longer-term strategies for protecting patients' rights and for sharing data warrant continuing discussion because of the fast-changing environment that includes advancing technologies, data sharing initiatives by other countries, the broader movement favouring open science and open data, and new needs for health care and disease prevention. We recommend development of a mechanism for continuing monitoring of these developments, perhaps by means of an interdisciplinary platform or forum that would also help to raise public awareness of European achievements and problems in the area. We propose that academies could help to catalyse the start of this new function, mobilising researchers and research funders in broader discussion with policy-makers and other stakeholders. The voice of patients is also critically important

in these discussions. As one next step, the academy networks will use this report to bring the issues to the attention of research organisations and others across the EU/EEA and to academies worldwide.

9 | Further international discussion and coordination are needed

European politicians should also be more active to address relevant issues internationally through diplomatic channels. For example, problems caused by US intelligence legislation affecting research data are an issue that cannot be resolved within the EU. While this particular problem (core to Schrems II) is a large challenge, more generally the EU should seek to lead global discussion to encourage all countries to adopt appropriate privacy rules and discuss their option for regulatory reform to facilitate reciprocity in sharing data.

Appendix 1 | Working Group composition and timetable

The report was prepared by consultation with a Working Group of experts who discussed a range of issues:

- Approaches to protecting patient's rights during research, and the ethical obligation to make best use of the data.
- New opportunities to collect and use data to tackle health priorities.
- Development of the GDPR and its acknowledgement of the importance of research.
- Current limitations of the GDPR in facilitating international transfer of data.
- Understanding the wide range of benefits accrued from the sharing of personal data for health research, their linkage to SDGs, the need to continue pursuing health priorities, and an indication of what is now being lost in consequence of the recent impediments.
- Particular issues for urgent transfer of data highlighted by the COVID-19 pandemic.
- Exacerbation of current problems by the new status of the UK post-Brexit.
- Assessment of the current solutions for data transfer provided by the GDPR and the operational problems for public sector researchers in seeking to use these solutions.
- Options for improving the GDPR mechanisms and the importance of supplementary measures.
- The relevance of other key EU initiatives, such as the European Health Data Space and the European Open Science Cloud.
- The potential of PET options to mitigate data security and privacy risks.

The experts acted in an individual capacity and were nominated by member academies of ALLEA, EASAC and FEAM:

George Griffin (co-chair, UK)
Volker ter Meulen (co-chair, Germany)
Adelin Albert (Belgium)
Heidi Beate Bentzen (Norway)
Jan-Willem Boiten (The Netherlands)
John Danesh (UK)
Annette Grueters-Kieslich (Germany)
Katrin Kaarna (Estonia)
Külli Kingo (Estonia)
Christian Lovis (Switzerland)
Jose Pereira Miguel (Portugal)
Michael Parker (UK)
Johan Rung (Sweden)
Ursula Schmidt-Erfurth (Austria)
Lorenzo Simonato (Italy)
Indra Spiecker genannt Döhmann (Germany)
Krzysztof Tomasiewicz (Poland)
Giske Ursin (Norway)
Jaak Vilo (Estonia)
Rosa Castro (FEAM), Robin Fears (EASAC), Robert Vogt (ALLEA) (secretariat)

The project proposal was initially presented by the Norwegian Academy of Science and Letters to the Biosciences Steering Panel of EASAC in October 2019. The proposal was approved by the Councils of EASAC, FEAM and ALLEA in November–December 2019 and announced on the academies' websites in April–May 2020, together with a call for evidence.

The Working Group met by video conference in June 2020 and September 2020, together with Joos Vandewalle (ALLEA, The Netherlands).

The scope of the project was discussed with Alisa Vekeman of the Directorate-General for Justice and Consumers in June 2020.

In addition to the Working Group meetings, evidence was gathered in a virtual workshop organised by the FEAM Forum in October 2020 (see Appendix 4).

The draft report was peer-reviewed by academy-nominated experts from December 2020 to February 2021 and the consensus report endorsed by member academies.

ALLEA, EASAC and FEAM thank all who contributed to preparing and reviewing the text.

Appendix 2 | Value of sharing personal data for health research

1 | Examples discussed in previous work of FEAM and the Wellcome Trust (2012)

- Multi-national European study providing evidence that prenatal treatment of toxoplasmosis had no effect on mother to child transmission of infection, leading to important policy changes on neonatal screening.
- Finnish registry data study on improving treatment options for schizophrenia identified opportunities for reducing mortality attributed to suicidal deaths.
- Multi-national European registry study on pre-term birth as a risk factor for developing high blood pressure as adults, leading to better understanding of pathology and new opportunities for improved control.
- Multi-national European study investigating link between diabetes treatment and the occurrence of cancer, likely to provide robust evidence for new treatment recommendations and improved patient care.
- Previous legislation in Germany requiring informed consent resulted in temporary breakdown of oldest cancer registry worldwide. Continuing challenges to disease registries might also have significant effects on capacities for disease monitoring and research.
- UK series of investigations to understand the link between smoking and lung cancer have saved millions of lives but would have been difficult to conduct if consent was required from data subjects.

2 | Examples of using big data for improved research and treatments, cited by European Commission[51]

Examples discussed of EU-funded multi-national projects on asthma, hearing loss, child obesity, Parkinson disease and work-related stress.

3 | Examples from recent peer-reviewed literature for the value of combining data sets in health research[52]

- National (The Netherlands) integration of information from epidemiological, pharmacological, genetic and gut microbiome databases as a tool for pharmaceutical research, to improve drug efficacy, safety and repurposing (Liu et al. 2020).

51| See https://ec.europa.eu/digital-single-market/en/managing-health-data.

52| Other examples are discussed in previous academy network reports. For example, the health value created by combining different data sets from epidemiological and environmental sources is described by EASAC (2019).

- Longitudinal US study to explore association of lead-exposure risk and family income with childhood brain outcomes (Marshall et al. 2020).
- Israeli study based on national electronic health records to predict gestational diabetes (Artzi et al. 2020). This study used a machine-learning predictor trained and validated on records from a single country; the authors noted that the applicability to other populations requires international comparison.

4 | Examples from the Scientific Panel for Health conference[53]: EU-funded (Horizon 2020) multi-national research demonstrating value for patients

- **www.rtcure.com:** public–private partnership under IMI2 to prevent rheumatoid arthritis and its progression. EU countries plus Australia.
- **http://chrodis.eu:** implementing good practices for chronic diseases.
- **www.secure-h2020.eu:** secondary prevention of cardiovascular disease in the elderly.
- **www.icpermed.eu:** international consortium for personalised medicine. Platform to initiate and support communication and exchange on personalised medicine research, EU countries plus Brazil, Canada, Iran, Turkey and Israel.

53| "Recommendations for health research in Europe – design for impact", Final Annual Conference of Scientific Panel for Health, July 2020.

5 | Collaborative research projects that have been affected by the rules governing the transfer of health data for research outside the EEA[54]

- **University College Dublin:** researchers at University College Dublin had to split one research study into two separate parts, one within the EU and one outside the EU, to comply with GDPR standards. This allowed the research team to maintain separate data sets without any cross-border transfers, comparing only the meta-analyses of each study. While the researchers could conduct their study and comply with law, modifications such as this may increase costs, affect statistical analysis and sample size, and increase the possibility of inaccuracy.[55]
- **For new projects in the National Cancer Institute Cohort Consortium,** European investigators have suggested that analyses either take place in Europe and then results are submitted to the USA for their combination in a meta-analysis, or that all the analyses must take place in Europe[56].
- **The International Genomics of Alzheimer's Consortium and the Alzheimer's Disease Sequencing Project:** researchers in the two organisations have not been able to pool and process data sets with personal data,

54| Examples included in this section are mainly based on notes from an interview with US researcher Dr Tetsuo Ashizawa Houston Methodist, Weill Cornell Medical College (6 August 2020), and personal communication from Dr Robert Eiss, NIH's Fogarty International Center (4 September 2020) as well as from the publications referred below.

55|http://iscintelligence.com/archivos_subidos/input_paper_on_gdpr_challenges_for_research-77623379-v15.pdf.

56| Personal communication from Dr. Giske Ursin, Director of the Cancer Registry of Norway.

since the first organisation is based in the EU and the second is conducted by the University of Pennsylvania, Perelman School of Medicine. Processing these data sets together would have allowed a more streamlined research process for novel drug targets for Alzheimer disease.

- **Clinical Trial Readiness for SCA1 and SCA3 (READISCA), National Institute of Neurological Disorder and Stroke (NINDS) Grant No. U01 NS104326:** The READISCA project funded by the NIH's NINDS aims to strengthen clinical trial readiness for spinocerebellar ataxia type 1 and type 3. Since these ataxias are rare diseases, the study relies on data from participants at 20 sites in the USA and two sites in Europe, specifically the Institut du Cerveau et de la Moëlle Epinière (ICM) in France and the University Hospital Bonn in Germany. The NIH believes that data sharing is essential for expedited translation of research, results into knowledge, products, and procedures to improve human health, so it has a data sharing policy in place that expects the timely release and sharing of final research data from NIH-supported studies for use by other researchers[57] including academic and industry researchers. The READISCA investigators, while working on the subaward with the European Institutions, could not sign a number of clauses required by the GDPR, including those specifying indemnification of collaborators, ownership of generated data, definition of personal data, designation of "Data Controller" and "Data Processor" associated with the management and control of the personal data, auditing of data systems by a foreign entity, and submitting to the jurisdiction of foreign courts. The issue was further complicated by the plan to use data of study participants who had been in observational clinical studies that took place in 2009-2012 with similar clinical outcome assessments. The parties were able to come to a compromise where the research team applied their best efforts to re-inform participants already enrolled in the studies about data sharing between the US and Europe. While they were eventually able to continue research by restricting the data sharing to investigators of neurodegenerative diseases, their work was delayed by almost two years with long legal discussions. About five other studies in the Clinical Trials Readiness project were similarly affected[58].

- **Statens Serum Institut (SSI):** SSI in Copenhagen provides data and biosamples from the Danish National Birth Cohort to the NIH to identify risk factors for type 2 diabetes and related co-morbidity. With the enactment of the GDPR, SSI claimed that these samples were no longer legally allowed in the USA and asked the NIH to return the data, unless they could either sign the EU-US Privacy Shield or sign the GDPR Article 46 SCCs. Since the NIH is not an organisation regulated by the Federal Trade Commission and the Department of Transportation, it could not sign the Privacy Shield. It was also not allowed to sign the SCCs because of conflicts with US law. The NIH proposed negotiations with SSI and the Danish

57| See https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html.

58| Notes from interview with Dr. Ashizawa (6 August 2020), and personal communication from Robert Eiss, 4 September 2020.

Ministry of Health, but SSI is demanding a recall of samples and has refused to provide additional data required under the original contract, preventing the research from continuing. SSI, which also houses the Danish National Biobank, has also frozen data streams to the WHO's International Agency for Research on Cancer in France (Rabesandratana 2019; Peloquin et al. 2020).

- **US National Cancer Institute Cohort Consortium:** multiple research projects, including those funded by the US NIHA where European institutions were supposed to contribute data, have been affected (Ursin et al. 2019a).

- **The International Agency for Research on Cancer at the WHO:** multiple studies have been affected at global level[59].

- **NIH:** In general, the GDPR has stalled at least 40 clinical and observational studies at the NIH on risk factors and exposures for cancer (Eiss 2020).

59 | https://edpb.europa.eu/sites/edpb/files/webform/public_consultation_reply/2020.05.14_letter_to_edpb_chair_with_un_comments_on_guidelines_2-2020.pdf.

Appendix 3 | PETs: types, potential uses and limitations

While a universal taxonomy does not exist for PETs, several studies have attempted to group PETs into different categories according to their technical contributions, functionalities or other criteria. Among them, the European Union Agency for Network and Information Security (ENISA) identified four categories: (1) secure messaging; (2) virtual private networks; (3) anonymising networks; and (4) anti-tracking tools for online browsing (ENISA 2016). Likewise, the Office of the Privacy Commissioner of Canada classified PETs according to whether they provide support or solve problems related to the following issues: informed consent, data minimisation, data tracking, anonymity, control, negotiate terms and conditions, technical enforcement, remote audit of enforcement and use of legal rights (Office of the Privacy Commissioner of Canada 2017).

Among the PETs offering different degrees of solutions to address these issues with potential implications for health research are the following.

- Homomorphic encryption allows data to be encrypted before it is shared, so that other parties can analyse the data without deciphering it. There are several possibilities for homomorphic encryption; for instance, fully homomorphic encryption enables data to be stored in encrypted code in the cloud while still allowing calculations to be performed on it (Gentry 2009). Guidance from the EDPB on whether this would still constitute processing of personal data would be welcome. On the downside, homomorphic encryption is computationally expensive and less productive than working with unencrypted data. Particular types of homographic encryption also face some limitations. For instance, fully homomorphic encryption only works for very simple calculations (simple statistical parameters), because the overhead and complexity of the calculations evolves very rapidly. However, these concerns are being addressed by ongoing research.
- Differential privacy consists of the addition of noise to an analytical system so that it is impossible to reverse-engineer the individual inputs. Limitations of differential privacy are related to the perturbation of the dataset; for instance, data manipulation can degrade the data. Overall, the widespread application of differential privacy, especially in health-related areas such as medical imaging, seems to need further research (Kaissis et al. 2020).
- Federated analysis allows parties to share the insights from their analysis without sharing the data itself. They have "arguably become the most widely used next-generation privacy preservation technique, both in industry and medical AI applications" (Kaissis et al. 2020). While federated learning/analysis provides

a good response to data governance issues, it poses a problem of data security and trust. In addition, decentralisation can make the curation of data and the maintenance of the integrity and quality of results more difficult; in a federated approach one cannot scrutinise the data, and therefore one runs the risk of a trade-off in the quality of the data and the analysis. Moreover, unless it is used in combination with other techniques, such as encryption of data, this technique does not completely solve security and privacy issues.

- Confidential computing "uses hardware-based techniques to isolate data, specific functions, or an entire application from the operating system, hypervisor or virtual machine manager, and other privileged processes. Data is stored in the trusted execution environment (TEE), where it's impossible to view the data or operations performed on it from outside, even with a debugger".[60] Confidential computing could facilitate the combination of data sets from multiple organisations for analysis without the need to provide access to each other's data. However, confidential computing is a very specific solution for specific problems, and therefore may become very expensive and inflexible when used more widely.
- Secure multi-party computation foresees data analysis to be spread across multiple parties in a way that no individual party is able to see the complete set of inputs. While serious progress has been made in the area of cryptographic techniques in recent years, and as a result joint calculation on data that remain stored locally and protected (multi-party computation) can be performed, one limitation is the need for continuous data transfer between parties, and the need for continuous online availability. Also, the reliability/redundancy and scalability to more than a small number of parties can limit its application; computational considerations are an additional concern, with current techniques being still two to three orders of magnitude slower than when all data are put in the cloud.
- Blockchain is an open, distributed ledger that can record transactions between several parties efficiently and in verifiable and permanent way; although it is not considered a PET, it can be made more privacy-preserving using PETs (Royal Society 2019). Nonetheless, the process is computationally expensive, and can be particularly relevant in systems where no party can act as a trusted party.

Some limitations of the use of PETs include the following.

- The risk of losing utility and accuracy in the data used. For instance, adding noise to a dataset (with differential privacy) can lead to losing some useful information, which can in turn diminish the accuracy of data.
- The risk of losing the ability to scrutinise the data before using it, namely running the risk of "garbage in, garbage out". While some techniques might deal with this problem to some extent, this activity certainly becomes more complex.

60 | https://spectrum.ieee.org/computing/hardware/what-is-confidential-computing.

- The use of encrypted data for computation (e.g. with homomorphic encryption and secure multi-party computation) may also entail the use of significant computation resources (e.g. time, computing power). Studies have indeed described a trade-off between utility and privacy with the use of PETs. However, further evolution of these technologies could affect the utility trade-offs of implementing more PETs in the future.
- The financial costs of using PETs are another important limitation. These costs depend on many aspects, including the type of data and PET used, and the scale of the project.
- Different PETs are also at different levels of technology readiness (a term used to describe whether a PET is close or not to being used or deployed in practice, and on which scale). Therefore, while many of these techniques are presented as the "silver bullet" to solve some privacy-related issues, often they are themselves at a research stage.

The above-mentioned technologies, along with a combination of other approaches, have been used in several projects using health data for research purposes.

- Homomorphic encryption. A genome-wide association study – a study aimed at identifying genetic variants associated with a trait – was performed by using a framework based on homomorphic encryption. This allowed the researchers to perform accurate genome-wide association studies for a real dataset of more than 25,000 individuals, while maintaining all individual data encrypted (Blatt et al. 2020). Also, in the Privitar de-identification project, the UK National Health Service (NHS) used consistent tokenisation and partially homomorphic encryption to de-identify health data from individuals.[61]
- Data standardisation and federated analysis. The Observational Medical Outcomes Partnership (OMOP) common data model allows for systematic analysis of disparate observational databases by transforming data within different databases into a common format and representation.[62] The European Health Data and Evidence Network (EHDEN) project built a large-scale, sustainable, federated network of standardised data sources using OMOP (a data standardisation tool) for the harmonisation of anonymised health records.
- Federated analysis. Personal Health Train – an initiative of a large Dutch coalition led by the Dutch Techcentre for Life Sciences[63]– aims to connect distributed health data and create value by increasing the use of existing health data for citizens, health care and scientific research. The key concept is to bring algorithms to the data where they happen to be rather than

61| Stuart Gunson, De-Identification Project Manager, Data Processing Services Programme, NHS Digital.

62| https://www.ohdsi.org/data-standardization/the-common-data-model/. See also Bergquist, Tim, and Pascal Brandt. "Prometheus: Differential Privacy in the OMOP CDM." (2018), https://courses.cs.washington.edu/courses/cse544/18wi/project/examples-successful-projects/psbrandt.pdf, describing the use of differential privacy in the OMOP common data model. Observational Health Data Sciences and Informatics (OHDSI), www.ohdsi.org, is a multi-stakeholder, interdisciplinary collaboration to bring out value in health data through large-scale analytics; it supports work on health data quality standardisation and identifies software opportunities to share/access data. OHDSI Europe is coordinated by Erasmus University Medical Centre, Rotterdam: www.ohdsi-europe.org.

63| www.dtls.nl/fair-data/personal-health-train.

bringing all data to a central place, giving controlled access to heterogeneous data sources while ensuring privacy protection (according to FAIR principles: findable, accessible, interoperable, reusable).

- Federated analysis. OpenSAFELY provides a secure software interface that allows detailed pseudonymised primary care patient records to be analysed in near real-time where they reside, hosted within the electronic health vendor's highly secure data centre. The purpose is to minimise the re-identification risks when data are transported off-site (Williamson et al. 2020). While this initiative provided a suitable solution for sharing data in a manner that could accommodate the urgency of COVID-19 with the need to protect personal data, the analysis of these data has been performed in pseudonymised primary care records, which carried their own limitations, and has only been shared within one country (UK).
- Blockchain. MyHealthMyData (MHMD). An EU-funded project (H2020), MyHealthMyData (MHMD) aims to change the way sensitive data are shared and to be the first open biomedical information network centred on the connection between organisations and individuals, encouraging hospitals to start making anonymised data available for open research.[64]

64| http://www.myhealthmydata.eu/.

Appendix 4 | Short summary from the FEAM European Biomedical Policy Forum

The FEAM Forum organised a webinar on "International transfer of health data" on 16 October 2020: see https://www.feam.eu/events/webinar-international-transfer-of-health-data-16-october-2020/. The FEAM European Biomedical Policy Forum is a platform for discussion of key policy issues, bringing together representatives from academia, research charities, industry, European and national trade associations and professional bodies, and patient and consumer groups.

A summary of this event has been published by FEAM[65]; a very brief account of the scope is indicated below and points of detail have been used to inform chapters 1–5.

Giske Ursin (Cancer Registry of Norway) discussed the importance of sharing pseudonymised health data outside the EU/EEA. GDPR implementation has presented obstacles to sharing data for public sector research and, with international organisations such as the WHO (International Agency for Research on Cancer) for public health purposes. The GDPR provides a layered approach to protecting data but there is urgent need for a workable mechanism for public institutions that do not conflict with other countries' laws, and for practical guidelines on international transfer.

Alisa Wakeman (European Commission, DG Justice) agreed that international data transfer is very important, the European Commission has recognised the concerns expressed by the research community and acknowledges that more EDPB guidance is needed. The SCCs are being updated to cover more scenarios and to take account of the implications of the recent ECJ (Schrems II) judgement. Further input of evidence and advice from the research community is welcome.

Laura Drechsler (researcher at Vrije Universiteit Brussel, Belgium) observed that the GDPR regime is complicated, and unclear in its scope on data transfer. Options for transfer mechanisms depend on a high standard of privacy protection in recipient countries and it can be an arduous task for EU/EEA research institutions to assess the level of protection offered by the recipient. Health data are regarded by the GDPR as a special category in need of protection but Member States can define their own safeguards, resulting in a complex landscape of regulation.

Brendan Barnes (European Federation of Pharmaceutical Industries and Associations) explained how the European research-

65| See https://www.feam.eu/wp-content/uploads/ITHD-Summary-report-5-Nov-2020-FINAL.pdf.

based pharmaceutical industry expressed considerable concern that the Schrems II judgement had put private sector data flows in doubt. The judgement had reflected EU concerns about US mass surveillance for national security purposes rather than any issue for the health sector, but the consequences for the flow of health data are serious. Health data, for example from clinical trials, regulatory submissions, "real-world" studies, are vitally important for public health and are already regulated by multiple safeguards such as research ethics approval and good clinical practice frameworks, as well as contractual standards.

Carlos Luis Parra Calderon (European Federation of Medical Informatics) described experience, including recently with COVID-19, in transferring data within Europe and with the USA. Advances in informatics can help to increase the security of data, and the GDPR requirements can help to stimulate advances in technologies enabling data sharing. Such advances may overcome some of the present limitations of PETs, for example for detailed analysis of federated data.

Gozde Susuzlu (European Patients' Forum) highlighted how patients are aware of the issues for security of data and how technology advances may help. In the experience of the European Patients' Forum, patients agree with the view that sharing of data across international borders is important.

Panel discussion provided further detail on many of these points; for example:

- The need to document examples of the value of sharing health data internationally (and where examples of good practice can help other researchers).
- The problems created by the impact of the GDPR on health data sharing. Because there has also recently been increasing citizen awareness of their privacy rights and concern about international differences in protecting privacy, there are differing perspectives on the extent to which it has been GDPR implementation that has created new obstacles.
- Timetable for European Commission action: progress is expected soon on providing draft guidelines and examples to develop concrete tools. There is now an opportunity for the research community to advise on examples of good practice already in place.
- There is potential for exploring other GDPR data transfer mechanisms, especially the public interest derogation, although if this were to be used systematically for transfers then individuals would not benefit from current protections.

Glossary

Anonymous information (Recital 26, GDPR) "information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable". The principles of data protection do not apply to anonymous information.

Data concerning health (Article 4, GDPR) "means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status".

European Health Data Space. The creation of a common European Health Data Space that "will promote better exchange and access to different types of health data (electronic health records, genomics data, data from patient registries etc.), not only to support healthcare delivery (so-called primary use of data) but also for health research and health policy making purposes (so-called secondary use of data)" is one of the priorities of the EU Commission 2019–2025. The exact definition of the European Health Data Space is still work in progress as the EU Commission is currently engaged in preparatory work with Member States to better define this proposal[66].

International transfer. Under the GDPR, any transfer of personal data that are undergoing processing or are intended for processing after transfer to a third country (outside the EU/EEA) or to an international organisation, and that are therefore subject to the conditions laid down by chapter V of the GDPR. Provision of remote access is also considered transfer.

Pseudonymisation (Article 4, GDPR) "means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".

Personal data (Article 4, GDPR) "means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, 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"

Public authority or body. According to the EDPB Guidelines,[67] which follow a broad approach covering both public bodies in third

66| https://ec.europa.eu/health/ehealth/dataspace_en.

67| Guidelines 2/2020 on Articles 46(2)(a) and 46(3)(b) of Regulation 2016/679 for transfers of personal data between EEA and non-EEA public authorities and bodies.

countries and international organisations, the definition of public bodies in third countries should be determined by domestic law and therefore could include public bodies include government authorities at different levels (e.g. national, regional and local authorities), as well as other bodies governed by public law (e.g. executive agencies).

Public research sector. It can be defined as those who perform or finance research and experimental development for the government, higher education institutions, public research bodies or other non-profit institutions. See Bentzen (2020) for a discussion of the elements the ECJ used to define scientific research.

Scientific research. While the GDPR does not include a definition of scientific research, Recital 159 establishes that "processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research. In addition, it should take into account the Union's objective under Article 179(1) TFEU of achieving a European Research Area. Scientific research purposes should also include studies conducted in the public interest in the area of public health. To meet the specificities of processing personal data for scientific research purposes, specific conditions should apply in particular as regards the publication or otherwise disclosure of personal data in the context of scientific research purposes. If the result of scientific research in particular in the health context gives reason for further measures in the interest of the data subject, the general rules of this Regulation should apply in view of those measures".

Abbreviations

ALLEA	European Federation of Academies of Sciences and Humanities
DG	Directorate-General
EASAC	European Academies' Science Advisory Council
ECJ	European Court of Justice
EDPB	European Data Protection Board
EEA	European Economic Area
ENISA	European Union Agency for Network and Information Security
EPRS	European Parliamentary Research Service
EU	European Union
FEAM	Federation of European Academies of Medicine
GDPR	General Data Protection Regulation
NHS	National Health Service (UK)
NIH	National Institutes of Health (USA)
OMOP	Observational Medical Outcomes Partnership
PET	Privacy-enhancing technology
SCC	Standard contractual clause
SDG	Sustainable Development Goal
SSI	Statens Serum Institut (Denmark)
WHO	World Health Organization

References

ALLEA and Royal Society. "*Flourishing in a Data-Enabled Society.*" ALLEA Discussion Paper 4, 2019.

Amann, RI, Baichoo, S, Blencowe, BJ, et al. "Toward unrestricted use of public genomic data." *Science* 363 (2019): 350–352.

Artzi, NS, Shilo, S, Hadar, E, et al. "Prediction of gestational diabetes based on nationwide electronic health records." *Nature Medicine* 26 (2020): 71–76.

Bentzen, HB. In the name of scientific achievement. How to assess what constitutes "scientific research" in the GDPR to protect data subjects and democracy. In *Disinformation and Digital Media as a Challenge for Democracy*. Ed. Terzis, G, Kloza, D, Kuzelewska, E, Trottier, D. European Integration and Democracy Series Volume 6. Intersentia, 2020. Chapter 18.

Bentzen, HB, and Hostmoelingen, N. "Balancing protection and free movement of personal data: the new European Union General Data Protection Regulation." *Annals of Internal Medicine* 170 (2019): 335–337.

Blatt, M, Gusev, A, Polyakov, Y, Goldwasser, S. "Secure large-scale genome-wide association studies using homomorphic encryption." *Proceedings of the National Academy of Science of the United States of America* 117.21 (2020): 11608–11613.

Brack, M, and Castillo, T "*Data Sharing for Public Health: Key Lessons from Other Sectors.*" Centre on Global Health Security, Chatham House, 2015.

Budin-Ljosne, I, Teare, HJA, Kaye, J, et al. "Dynamic consent: a potential solution to some of the challenges of modern biomedical research." *BMC Medical Ethics* 18 (2017): DOI:10.1186/s12910-016-0162-9.

Canova, C, Simonato, L, Barbiellini, C, et al. "A systematic review of case-identification algorithms for 18 conditions based on Italian healthcare administrative databases: a study protocol." *Epidemiologia e prevenzione* 43.4 (Suppl 2) (2019): 8–16.

Capurro, D, Cole, K, Echevarria, MI, et al. "The use of social networking sites for public health practice and research: a systematic review." *Journal of Medical Internet Research* 16 (2014): e79.

CIOMS in collaboration with the World Health Organization (WHO). "International Ethical Guidelines for Health-related Research Involving Humans." https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf. 2016.

Cleary, M. "The new digital healthcare ecosystem: looking outside to harness from within." *Value & Outcomes Spotlight* January/February (2020): 16–19.

Courbier, S, Dimond, R, and Bros-Facer, V. "Share and protect our health data: an evidence based approach to rare disease patients' perspectives on data sharing and data protection – quantitative survey and recommendations." *Orphanet Journal of Rare Diseases* 14 (2019): 175.

COVID-19 Clinical Research Coalition. "Global coalition to accelerate COVID-19 clinical research in resource-limited settings." *Lancet* (2019): https://doi.org/10.1016/S0140-6736(20)30798-4.

Cruz Rivera, S, Kyte, DG, Lee Aiyegbusi, O, Keeley, TJ, and Calvert, MJ. "Assessing the impact of healthcare research: a systematic review of methodological frameworks." *PLOS Medicine* 14 (2017): e1002370.

EASAC. "*The Imperative of Climate Action to Protect Human Health in Europe.*" Policy Report 38, 2019.

Eiss, R. "Confusion over data-privacy law stalls scientific progress." *Nature* 584 (2020): 498.

EPRS. "*How the General Data Protection Regulation changes the Rules for Scientific Research.*" STOA PE 634.447, 2020.

ENISA. "*PETs Controls Matrix - A Systematic Approach for Assessing Online and Mobile Privacy Tools.*" 20 December 2016.

European Commission. "*Data protection as a pillar of citizen's empowerment and the EU's approach to the digital transition – two years of application of the General Data Protection Regulation.*" COM (2020) 264 final, 2020.

EU Scientific Panel for Health. "*Better Research for Better Health. A Vision For Health and Biomedical Research from the Scientific Panel for Health.*" https://ec.europa.eu/programmes/horizon2020/en/h2020–section/scientific-panel-health-sph. 2016.

FEAM European Biomedical Policy Forum. "*Use of Data in Cross-Border Biomedical Research: What Are the Challenges Ahead for Europe?*". 2018.

FEAM European Biomedical Policy Forum. "*Artificial Intelligence in Healthcare: is Europe Ready?*" https://www.feam.eu/wp-content/uploads/AI-Summary-report-15-Apr-2019-FV.pdf. 2019.

FEAM European Biomedical Policy Forum. "*International Transfer of Health Data*" https://www.feam.eu/wp-content/uploads/ITHD-Summary-report-5-Nov-2020-FINAL.pdf. 2020.

FEAM and Wellcome Trust. "*Realising the Societal Benefits of Health Research through the Data Protection Regulation.*" www.feam.eu/wp-content/uploads/FEAMWTMEPbriefingonDPRamendments_amdts-1.pdf. 2012.

FEAM, EASAC et al. "*Protecting Health and Scientific Research in the Data Protection Regulation. Position of Non-Commercial Research Organisations and Academics.*" www.feam.eu/wp-content/uploads/LIBEreportJointStatementFebruary2014–1.pdf. 2014.

FEAM, EASAC et al. "*Ensuring a Healthy Future for Scientific Research through the Data Protection Regulation. Position of Academic, Patient and Non-Commercial Research Organisations.*" www.feam.eu/wp-content/uploads/Data_Protection_jointstatement_July_2015-1.pdf. 2015.

Fears, R, Brand, R, Frackiowiak, R, Pastoret, P-P, Souhami, R, and Thompson, B. "Data protection regulation and the promotion of health research: getting the balance right." *Quarterly Journal of Medicine* 107 (2014): 3–5.

Gentry, C. "*A Fully Homomorphic Encryption Scheme.*" PhD thesis, Stanford University. 2009.

Gesundheit Österreich Forschungs- und Planungs GmbH. "*Study on Big Data in Public Health, Telemedicine and Healthcare.*" https://ec.europa.eu/health/files/ehealth/docs/bigdata_report_en.pdf. 2016.

G-Science Academies. "*Statement: Digital health and the learning health system.*" 2020.

Hecker, S, Bonner, R, Haklay, M, et al. "Innovation in citizen science – perspectives on science-policy advances." *Citizen Science Theory and Practice* 3 (2018): http://doi.org/10.5334/cstp.114.

Hess, C, and Ostrom, E. *Understanding Knowledge as a Commons. From Theory to Practice*. MIT Press, 2011.

Hripsak, G, Duke, JD, Shah, NH, et al. "Observational health data sciences and informatics (OHDSI): opportunities for observational researchers." *Studies in Health Technology and Informatics* 216 (2015): 574–578.

Kaissis, GA, Makowski, MR, Rückert, D, and Braren, RF. "Secure, privacy-preserving and federated machine learning in medical imaging." *Nature Machine Intelligence* 2 (2020): 305–311.

Kalkman, S, van Delden, J, Banerjee, A, et al. "Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the empirical evidence." *Journal of Medical Ethics* (2019a): https://doi.org/10.1136/medethics-2019-105651.

Kalkman, S, Mostert, M, Gerlinger, C, et al. "Responsible data sharing in international health research: a systematic review of principles and norms." *BMC Medical Ethics* (2019b): https://doi.org/10.1186/s12910-019-0359-9.

Knoppers, B. "Broaden human-rights focus for health data under GDPR." *Nature* 558.7709 (2018): 189–190.

Knottnerus, JA. "Research data as a global public good." *Journal of Clinical Epidemiology* 70 (2015): 270–271.

Liu, J, Lahousse, L, van Duijn, CM, et al. "Integration of epidemiologic, pharmacologic. Genetic and gut microbiome data in a drug-metabolite atlas." *Nature Medicine* 26 (2020): 110–117.

Marshall, AT, Betts, S, Kan, EC, et al. "Association of lead-exposure risk and family income with childhood brain outcomes." *Nature Medicine* 26 (2020): 91–97.

Mascalzoni, D, Bentzen, HB, Budin-Ljosne, I, et al. "Are requirements to deposit data in research repositories compatible with the European Union's General Data Protection Regulation?" *Annals of Internal Medicine* 170 (2019): 332–334.

Middleton, A, Milne, R, Almarri, MA, et al. "Global public perceptions of genomic data sharing: what shapes the willingness to donate DNA and health data?" *American Journal of Human Genetics* 107 (2020): 743–752.

Mitchell, C, Ordish, J, Johnson, E, Brigden, T, Hall, A. "*The GDPR and Genomic Data: The Impact of The GDPR and DPA 2018 on Genomic Healthcare and Research.*" PHG Foundation, 2020.

Moorthy, V, Restrepo, AMH, Preziosi, M-P, and Swaminathan, S. "Data sharing for novel coronavirus (COVID-19)." *Bulletin of the World Health Organization* 98 (2020): 150.

Nicol, D, Eckstein, L, Bentzen, HB, et al. "Consent insufficient for data release." *Science* (2019): 10.1126/science.aax0892.

Nuffield Council on Bioethics. "*The Collection, Linking and Use of Data in Biomedical Research and Health Care: Ethical Issues.*" 2015.

Office of the Privacy Commissioner of Canada. "*Privacy Enhancing Technologies, A Review of Tools and Techniques.*" Report prepared by the Technology Analysis Division of the Office of the Privacy Commissioner of Canada, November 2017.

Ohmann, C, Banzi, R, Canham, S, et al. "Sharing and reuse of individual participant data from clinical trials: principles and recommendations." *BMJ Open* 7 (2017): e018647.

Peloquin, D, DiMaio, M, Bierer, B and Barnes, M "Disruptive and avoidable: GDPR challenges to secondary research uses of data." *European Journal of Human Genetics* 28 (2020): 697–705.

Phillips, M. "International data-sharing norms: from the OECD to the General Data Protection Regulation (GDPR)." *Human Genetics* 137 (2018): 575–582.

Phillips, M, Molnar-Gabor, F, Korbel, JO, et al. "Genomics: data sharing needs an international code of conduct." *Nature* 578 (2020): 31–33.

Rabesandratana, T. "Researchers sound alert on European data laws." *Science* 366 (2019): 936.

Royal Society. "*Protecting Privacy in Practice: The Current Use, Development and Limits of Privacy Enhancing Technologies in Data Analysis*". 2019.

Shabani, M, and Borry, P. "Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation." *European Journal of Human Genetics* 26 (2018): 149–156.

Shaffer, L. "WHO wants to bring order to health data." *Nature* Medicine 26 (2020): 2–3.

Sharon, T, and Lucivero, F. "Introduction to the Special Theme: the expansion of the health data ecosystem – rethinking data ethics and governance." *Big Data & Society* (July–December 2019): 1–5. www.journals.sagepub.com/doi/pdf/10.1177/2053951719852969

Shilo, S, Rossman, H, and Segal, E. "Axis of a revolution: challenges and promises of big data in healthcare." *Nature Medicine* 26 (2020): 29–38.

Sipido, KR, Antonanzas, F, Celis, J, et al. "Overcoming fragmentation of health research in Europe: lessons from COVID-19." *Lancet* 395 (2020): 1970–1971.

Slokenberga, S, Reichel, J, Niringiye, R, et al. "EU data transfer rules and African legal realities: is data exchange for biobank research realistic?" *International Data Privacy Law* 9 (2019): 30–48.

Starkbaum, J, and Felt, U. "Negotiating the reuse of health-data: research, big data, and the European General Data Protection Regulation." *Big Data & Society* (July-December 2019): 1–12. www.journals.sagepub.com/doi/pdf/10.1177/2053951719862594

Taylor, MJ, Wallace, SE, and Prictor, M. "United Kingdom: transfers of genomic data to third countries." *Human Genetics* 137 (2018): 637–645.

Ursin, G, Stenbeck, M, Chang-Claude, J, et al. "Data must be shared – also with researchers outside of Europe." *Lancet* 394 (2019a): 1902.

Ursin, G, Malila, N, Chang-Claude, J, et al. "Sharing data safely while preserving privacy." *Lancet* 394 (2019b): 1902–1903.

Verdonck, P, Hulle, MV, et al. "*Data Science and Healthcare.*" Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, Standpunten 48 b, 2018.

Walport, M, and Brest, P. "Sharing research data to improve public health." *Lancet* 377 (2011): 537–539.

Williamson EJ, Walker AJ, Bhaskaran K, et al. "Factors associated with COVID-19-related death using OpenSAFELY." *Nature* 584 (2020): 430–436.

About ALLEA

ALLEA is the European Federation of Academies of Sciences and Humanities, representing more than 50 academies from over 40 EU and non-EU countries. Since its foundation in 1994, ALLEA speaks out on behalf of its members on the European and international stages, promotes science as a global public good, and facilitates scientific collaboration across borders and disciplines. Jointly with its members, ALLEA seeks to improve the conditions for research, to provide the best independent and interdisciplinary science advice available, and to strengthen the role of science in society. In doing so, ALLEA channels the expertise of European academies for the benefit of the research community, decision-makers and the public.

About EASAC

EASAC is formed by the national science academies of the EU Member States, Norway, Switzerland and UK, to collaborate in giving advice to European policy-makers. EASAC provides a means for the collective voice of European science to be heard. Through EASAC, the academies work together to provide independent, expert, evidence-based advice about the scientific aspects of European policies to those who make or influence policy within the European institutions. Drawing on the memberships and networks of the academies, EASAC accesses the best of European science in carrying out its work. Its views are vigorously independent of commercial or political bias, and it is open and transparent in its processes. EASAC aims to deliver advice that is comprehensible, relevant and timely.

About FEAM

FEAM is the European umbrella group of national Academies of Medicine, Pharmacy and Veterinary Science, or national Academies via their medical division. It brings together under one umbrella 23 National Academies representing thousands among the best scientists in Europe. FEAM's mission is to promote cooperation between National Academies of Medicine and Medical Sections of Academies of Sciences in Europe; to provide a platform to formulate their collective voice on matters concerning human and animal medicine, biomedical research, education, and health with a European dimension; and to extend to the European authorities the advisory role that they exercise in their own countries on these matters.

All European Academies

Jaegerstrasse, 22/23
10117 Berlin | Germany
+49 (0)3 032 598 73 72
E-mail: secretariat@allea.org
Twitter: @ALLEA_academies
www.allea.org

European Academies
easac
Science Advisory Council

EASAC Secretariat
Deutsche Akademie der Naturforscher Leopoldina
German National Academy of Sciences
Jägerberg 1
06108 Halle (Saale) | Germany
+49 (0)3 454 723 98 33
E-mail: secretariat@easac.eu
Twitter: @EASACnews
www.easac.eu

FEAM
Federation of European Academies of Medicine

Rue d'Egmont, 13
1000 Brussels | Belgium
+32 (0)2 793 02 50
E-mail: info@feam.eu
Twitter: @FedEuroAcadMed
www.feam.eu