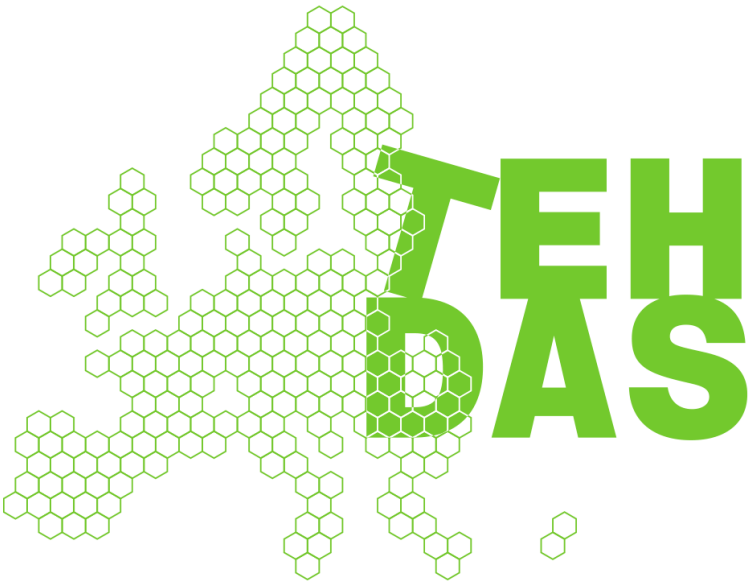


Towards European Health Data Space

Deliverable 5.1

Report on secondary use of health data through European case studies

28 February 2022

This project has been co-funded by the European Union's 3rd Health Programme (2014-2020) under Grant Agreement no 101035467.

0 Document info

0.1 Authors

Author	Partner
Linda Abboud	Sciensano, Belgium
Petronille Bogaert	Sciensano, Belgium
Sarion Bowers	Wellcome Sanger Institute, UK
Hayley Clissold	Wellcome Sanger Institute, UK
Shona Cosgrove	Sciensano, Belgium
Irène Kesisoglou	Sciensano, Belgium
Rosie Richards	NHS Confederation, UK
Catia Pinto	Servicos Partilhados do Ministerio de Saude (SPMS), Portugal
Miriam Saso	Sciensano, Belgium
Flávio Soares	Servicos Partilhados do Ministerio de Saude (SPMS), Portugal

0.2 Keywords

Keywords	TEHDAS, Joint Action, Health Data, European Health Data Space, Data Space, HP-JA-2020-1

Accepted in Project Steering Group on 22 February 2022. Updated on 30 August 2022.

Table of contents

1 Executive summary

The Joint Action Towards the European Health Data Space (TEHDAS JA) was launched under the Third Programme for EU Action in the Field of Health's 2020 Work Programme (3rd Health Programme). The overarching goal of TEHDAS is to support Member States and the Commission in developing and promoting concepts for exchanging health data for secondary purposes such as research, policy making, education and innovation across Europe.

TEHDAS is composed of eight work packages (WPs), exploring opportunities and creating proposals for data governance models and functions, as well as options for data quality management and data sharing infrastructures, sustainability and ethical models. The aim of WP 5 'Sharing Data for Health' is to provide recommendations for European countries on planning national legislation to enable cross-border exchange of health data. Within this scope, Task 5.1 aims to define and develop the evidence base for the secondary use of health data, from the perspective of data users (considered to be both researchers and policy makers).

An initial literature review identified barriers to cross-border sharing of health data for secondary use under the General Data Protection Regulation (GDPR). Based on the results, a data framework and a prioritisation analysis were developed to ascertain the most impactful and problematic barriers for data users. Thereafter, a survey was conducted to collect real-life examples and evidence on the impact of these barriers from data users, as well as to gather best practice examples to address them. The best practices identified in the case studies were refined and tested through consultations with data users. Based on that, policy options to address the top priority barriers to cross-border sharing of health data for secondary use were developed as set out in this report. All the options presented here are for consideration within the specific scope of the European Health Data Space (EHDS).

This report consolidates the results of the literature review, case studies and stakeholder consultations. The results showed that European data users experience a wide range of barriers to cross-border health data sharing for secondary use, mostly related to legal and data management issues caused by misalignment of interpretations and implementation (lack of semantic interoperability and differing interpretations of key terms). Regarding the real-life examples of the barriers identified (case studies), more than half were legal-related barriers, 30% were caused by data management, 13% were technical issues and 5% were trust-/transparency-related barriers. Data users also highlighted the real and serious impacts of the existing barriers that result in health data being underutilised for secondary use, reducing the benefits for all.

The options presented in this report specifically address the priority barriers experienced by data users and enable data sharing for secondary use for wider TEHDAS tasks and the European Commission to consider. Therefore, it is not within the scope of this report to suggest one preferred solution for each barrier.

Finally, the data user perspective and evidence base provided here will be used as the initial steps for different TEHDAS tasks ahead, as well as to support the decisions around the EHDS infrastructure, governance and legislation.

2 Introduction

The creation of a European Data Space is one of the priorities of the European Commission 2019-2025, including a specific data space for the health sector. A common European Health Data Space (EHDS) will promote better exchange and access to different types of health data, benefiting both healthcare delivery (also known as primary use of data), and research and policy making purposes (also known as secondary use of data).

The entire data system will be built on transparent foundations that fully protect citizens' data and reinforce the portability of their health data, in line with the General Data Protection Regulation (GDPR).

The **Joint Action Towards the European Health Data Space** (TEHDAS), helps EU Member States, associated countries and the European Commission to develop and promote the concepts necessary for the secondary use of health data, benefiting public health, research and innovation in Europe. The results of the TEHDAS project will provide elements to the European Commission's legislative proposal on the EHDS as well as support pan-European dialogue that will follow the proposal.

The project is divided into eight work packages led by organisations from different countries. The overall aim of Work Package 5 'Sharing Data for Health' is to develop options for governance models for the exchange and secondary use of health data between European countries, based on transparency, trust, citizen empowerment and for a common good. The work package will provide recommendations for European countries on planning national legislation to enable cross-border exchange and secondary use of health data.

Within this scope, Task 5.1 aims to define and develop the evidence base for the secondary use of health data, from the perspective of data users, to inform wider TEHDAS tasks and work packages as well as the development of the EHDS infrastructure, governance and legislation.

To achieve this aim, a three-step methodology was defined:

1. **Literature review:** A focused literature review was conducted to identify the barriers to cross-border sharing of health data for secondary use, for non-personal health data and personal health data under the General Data Protection Regulation (GDPR);

2. **Framework:** A framework was developed to provide an overview of the most impactful and problematic barriers to cross-border sharing of health data for data users in EU Member States and associated countries, based on a prioritisation of the literature review results;

3. **Case studies:** Case studies were developed in collaboration with data users to create the evidence for the cross-border sharing of health data for secondary use from their perspective in order to understand the impact of the barriers, identify best practices and suggestions to address them.

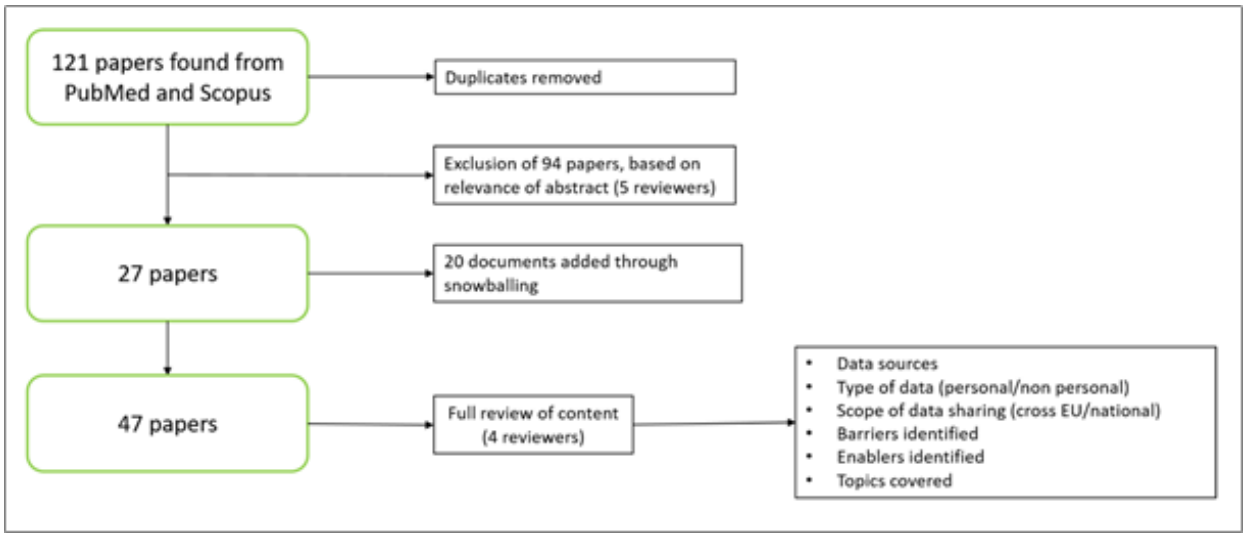
4. **Policy options:** The best practices from the case studies were analysed. They were then refined and tested through additional bilateral consultations with data users, to develop policy options.

This report consolidates the results of the literature review, case studies and stakeholder consultations. Each identified barrier is briefly presented including its impact as described by the data users. Thereafter, practical policy options for the European Commission, Member States and associated countries are presented to address barriers to the cross-border sharing of health data enabling secondary use. These options are based on the best practices provided by data users in the case studies and further refined through the stakeholder consultations. It is important to note that all the options presented in this report are developed for consideration within the specific scope of the EHDS.

3 Literature review

Firstly, a rapid literature review was carried out, in accordance with the Cochrane Rapid Reviews Methods Group guidance (Garritty et al., 2020)[1], between November 2020 and March 2021. The review was guided by the research question: 'What are the barriers and enablers to cross-border sharing of health data for secondary use, for non-personal health data and personal health data under the GDPR?'. A total of 121 papers were screened, and a final set of 47 papers were reviewed (see figure 1). Further details on the key words and inclusion criteria used for the rapid review can be found in Annex 1.

Figure 1: Presents the PRISMA flow diagram summarizing the rapid review process.



[1] Garritty, C., Hamel, C., Hersi, M. et al. Assessing how information is packaged in rapid reviews for policy-makers and other stakeholders: a cross-sectional study. Health Res Policy Sys 18, 112 (2020). https://doi.org/10.1186/s12961-020-00624-7

The rapid review produced a list of 89 barriers and 79 enablers for data sharing across Europe, which were compiled into a framework. As a starting point, the framework utilised the components of the 'Support tool to assess health information systems and develop and strengthen health information strategies' developed by the World Health Organisation (WHO) regional office for Europe[2]. Taking this into consideration, results of the literature review were categorised, and key patterns were identified and organised into the following four specific themes:

- **Data:** including data management, data quality, data interoperability, data monitoring and analysis
- **Infrastructure:** including the governance structure of the health data system and access to data
- **Legal:** including semantics, legal frameworks, and national interpretations of GDPR
- **Trust and transparency:** including political, social and organisational factors and citizens' engagement

4 Developing a framework: analysis of priority barriers

Based on the literature review findings, a framework was developed and the list of 89 barriers were refined to 20 barriers for further analysis based on frequency and impact calculations. The list included barriers representing the themes: data, legal, infrastructure, trust and transparency. Resources and ethical aspects were incorporated into the case study template as overarching themes. This list was further refined through a prioritisation exercise involving 18 countries' representatives in TEHDAS to ensure that resources and work to identify and develop could be tailored to the top priority barriers for data users.

The final list of 11 priority barriers is presented in table 1.

[2] World Health Organization. Regional Office for Europe. ((2021 Support tool of the WHO Support tool to strengthen health . information systems: guidance for health information system assessment and strategy development. Available at: https://www.euro.who.int/__data/assets/pdf_file/0011/278741/Support-tool-assess-HIS-en.pdf

Table 1: Final list of barriers as selected by participating TEHDAS countries

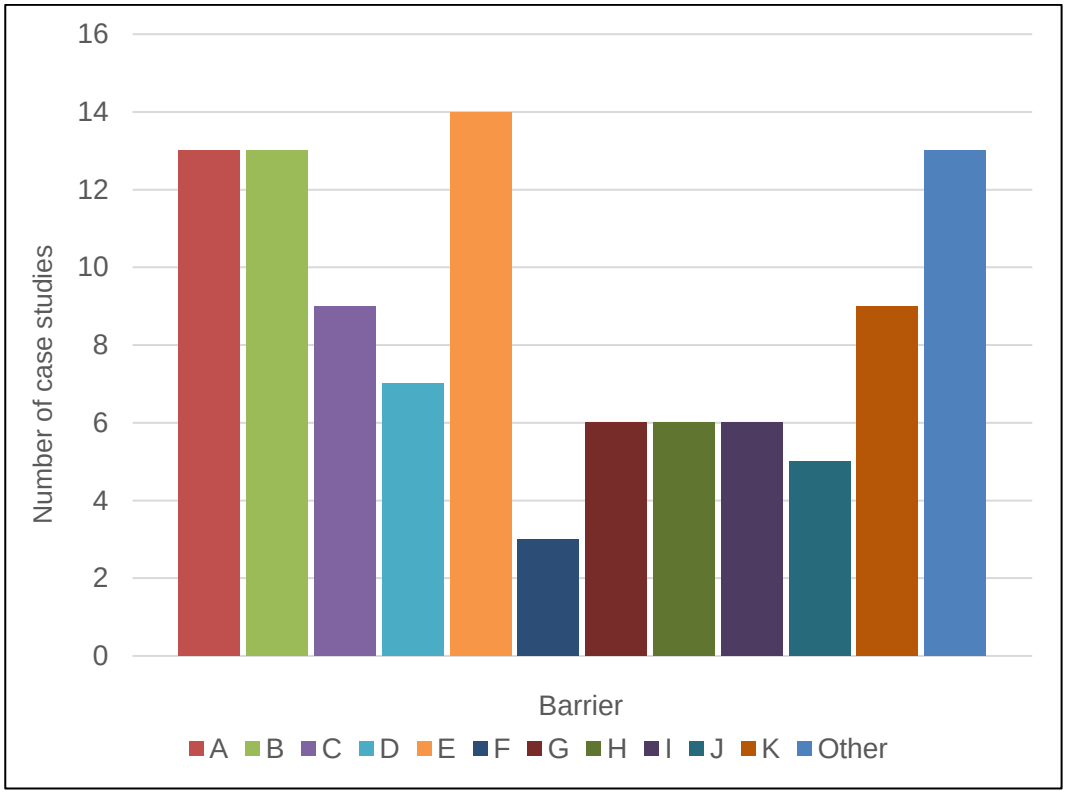
Rank	Barrier description	Theme
A	There are differences in governance and health data systems in Europe.	Infrastructure
B	A lack of a common European interpretation of what constitutes 'sufficient anonymisation' to transform personal data to non-personal data.	Legal
C	A lack of a common European interpretation of what constitutes 'pseudonymisation'.	Legal
D	A lack of a common European interpretation of what is and is not 'secondary use' of data.	Legal
E	European countries have national laws/rules on health and research data in addition to the GDPR.	Legal
F	European countries can set different derogations under the General Data Protection Regulation.	Legal
G	European countries have different preferences as to the choice of legal basis for processing personal data under the GDPR.	Legal
H	Health data is considered sensitive data e.g., special category data under the GDPR, and is treated differently from other types of data when it comes to health data ethics, management, and use.	Data
I	A lack of standardised data sharing agreements for products developed by private sector providers using public health data to facilitate safe data sharing and protect public investment.	Trust and Transparency
J	The use of different interoperability standards across Europe makes comparisons and sharing data and research results challenging.	Data
K	Poor data management procedures reduce the ability to reuse data.	Data

5 Case studies

The list of 11 priority barriers, as agreed by EU Member States' and associated countries' national representatives, formed the basis for the development of a survey to facilitate the collection of case studies from experts, institutes and/or projects within EU Member States or associated countries. The aim of the case studies was to substantiate the barriers, understand the impacts and further identify potential best practices and suggestions to address these barriers and improve data sharing between European countries. The case studies also aimed to explore the transfer purpose, requirements, and data type, focusing on scientific research and innovation and policy making for public health purposes to support T5.2, T5.4 and Work Packages 6 and 7.

A total of 23 European countries provided 113 case studies between April and August 2021. The following figures (1-3) provide an overview of the case studies submitted.

Figure 2: Number of entries per barrier



Legend:

- *A to K: see Table 1 for full descriptions of the barriers (lettered A to K)*
- *'Other': represents case studies for barriers that participants perceived as a priority but were not included in the predefined list of 11 priority barriers (lettered A to K)*

Figure 2 (Number of entries per barrier) shows the distribution of case study submissions per barrier. The barriers F and J had the lowest responses (≤ 5 responses). 'Other' represents case studies for barriers that participants perceived as a priority but were not included in the predefined list of 11 barriers.

Figure 3: Countries that contributed case studies

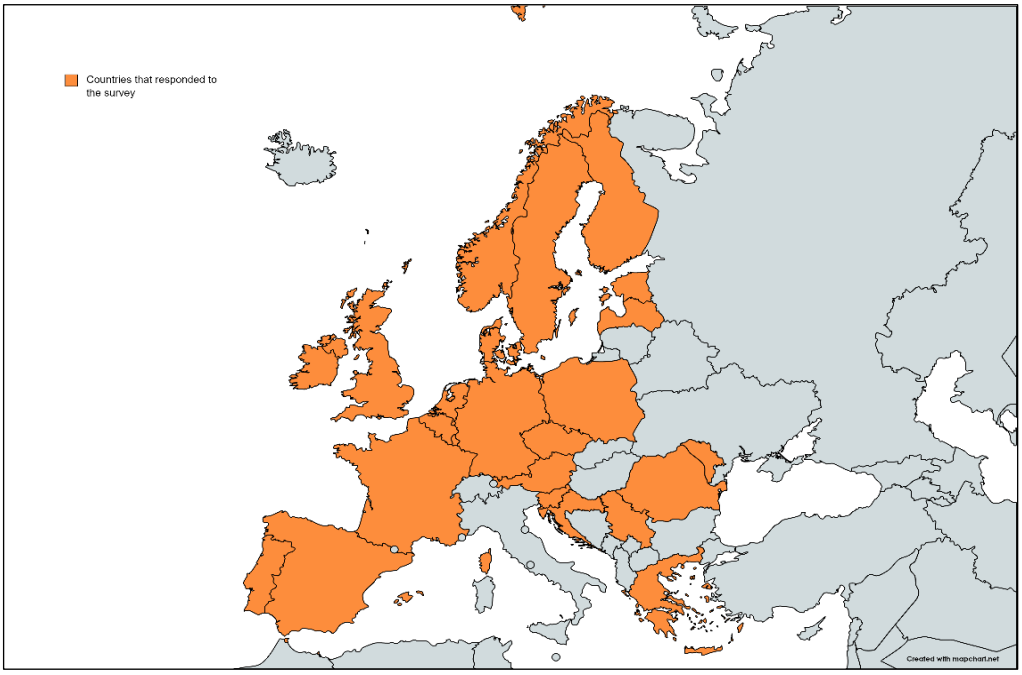
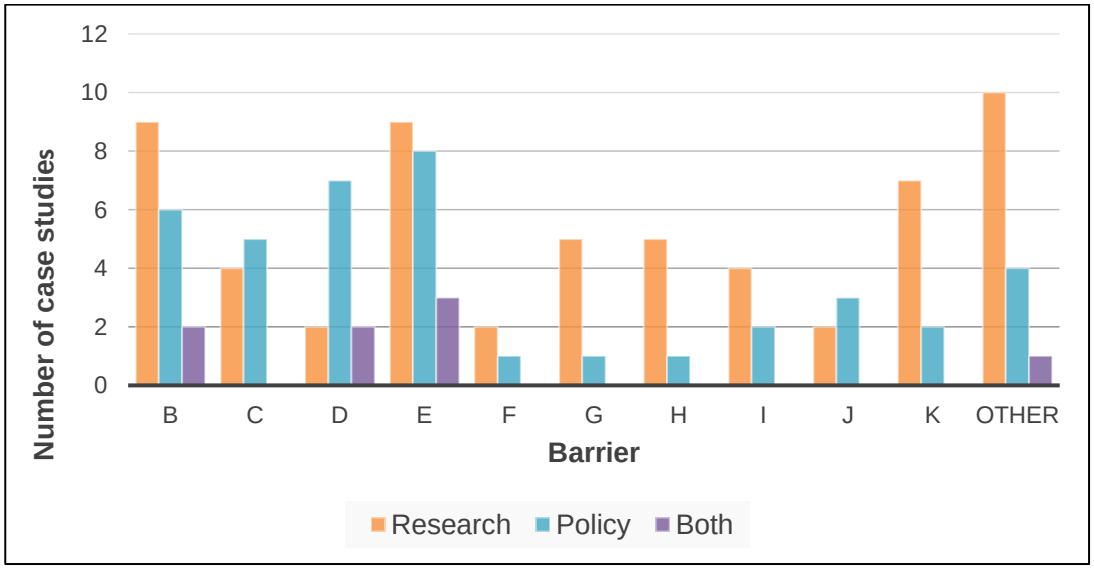


Figure 3 (Countries that contributed cases studies) provides the overview of the distribution of countries that responded to the survey. A total of 23 countries provided at least one case study, representing a wide coverage of European countries.

Figure 4: Case studies grouped according to sector



Legend:

- *B to K: see Table 1 for full descriptions of the barriers (lettered A to K)*
- *'Other': represents case studies for barriers that participants perceived as a priority but were not included in the predefined list of 11 priority barriers (lettered A to K)*
- *'Research': survey respondents who identified themselves as researchers*
- *'Policy': survey respondents who identified themselves as being involved in policymaking*
- *'Both': survey respondents who identified as both 'research' and 'policy' (i.e., selected both options in the survey)*

Figure 4 (Case studies grouped according to sector) shows the number of case studies submitted by users from the research and/or policy perspective. The 'both' category represents case studies where the contributor selected both 'researcher' and 'policymaker' in the job type field. It is important to note that barrier A (governance and health data management systems in Europe) is not included in Figure 4, as the inputs provide a description of the health data management system rather than an experience of the specific user (policymaker or researcher) with the barrier.

In general, each barrier received case studies from both policymakers and researchers, suggesting that all types of barriers are recognised by these different data users, which is an important finding. The best practices suggested by the case studies thus come from a variety of research and policy backgrounds. As a result, this report, generally refers to data users, meaning both researchers and policy-makers. Where relevant differences between data users have been identified, the specific term, either 'researcher' or 'policy maker' has been applied.

However, considering the responses received, the majority of case studies from policymakers related to barriers B to E, with much fewer policymakers submitting case studies on barriers F to K. Barriers B to E are legal barriers (see Table 1), which suggests that the policymakers who responded to the survey face more legal barriers to the secondary use of health data as opposed to other barriers (e.g., data-related barriers).

Conversely, limited conclusions can be made regarding the case studies provided by researchers. There is no clear trend of researchers providing more case studies for particular types of barriers, with legal, data and trust and transparency barriers being reported.

These conclusions should be interpreted bearing in mind the sample size of survey respondents who provided case studies.

6 Analysis of barriers for secondary use of health data

Each case study provides a real-life example of the impacts, issues and potential best practices and solutions to the barriers identified to cross-border data sharing. These were used to provide real-world evidence and stakeholder input to develop policy options to address barriers to cross-border health data sharing for secondary use. These options were further refined through consultations with 15 European experts who submitted case studies and whose recommendations on best practices were incorporated into the final policy options. The aim is to provide actionable options to the European Commission, and Member States and associated countries to address barriers to cross-border sharing of health data enabling secondary uses

Section 7 presents each barrier, its impacts as collected from the case studies, and the suggested list of policy options to take forward, based on best practices provided by data users. When reading the list, it is important to note that the options presented are not mutually exclusive and a combination approach could provide the most effective solution. In some cases, the options are complementary to each other, and may build on, or strengthen, each other. The options do not explicitly stipulate precisely who should be responsible for delivery, or the mechanisms for effecting these changes, as these would be decisions for the

European Commission. The options are listed starting from an EU wide approach, towards more national level action. The last option under every barrier is to 'do nothing' or maintain the status quo, which is included for completeness. For some of these options, adequate legislation, codes of conduct, resources and funding would need to be secured.

It is also important to note that all options in this document are considered within the specific scope of the EHDS. This is not explicitly re-stated in each option but is implicit, and thus the options should be interpreted within and limited to this scope.

6.1 Barrier A: There are differences in governance and health data systems in Europe

A cross-section of participating European countries was asked to map the governance and health data management systems in their countries in order to inform and support activities in T5.1, T.3 and Work Packages 4, 6 and 7. Ten countries (Austria, Belgium, Denmark, Estonia, Finland, Greece, Ireland, Moldova, Sweden, and the UK) provided a national-level snapshot. The research team ensured that the participating countries provide a good representation of countries across the EU and wider Europe.

Member States' and associated countries' case studies described national health data management models ranging from centralised models to decentralised and federated systems. Data users stressed that this divergent starting point would need to be taken into consideration in the development and implementation of digital health legislation as well as the underlying infrastructure for the European Health Data Space (EHDS). The descriptions provided by the countries are compiled in Annex 2, which presents the full health data management profiles for the participating countries.

Task 4.1 of TEHDAS is performing country visits to map the state of play of the health data management system in more depth and provide an overview to the European Commission in order to influence the development and implementation of the legislation on the EHDS. The plan is to map 12 EU countries by the end of 2022. The results will provide a more complete picture of this divergent starting point and complement the initial findings provided in Annex 2 of this deliverable.

The results of Barrier A helped build understanding of the health data systems across Europe in T5.2, T5.4, Work Package 6 and 7. The results were deliberately published in the milestone report MS5.1, 5.2) to provide early information, upon which these tasks and work packages have built further through their own milestone and deliverables.

6.2 Barrier B: A lack of a common European interpretation of what constitutes 'sufficient anonymisation' to transform personal data to non-personal data

Researchers and policymakers alike have identified a lack of guidance on anonymisation at national and international levels as a barrier to data sharing. Key issues include a lack of clarity between "absolute" and "relative" anonymisation, lack of guidance for specific types of

health data (e.g., medical images, genomic data, rare diseases), and how to define the parameters for re-identification.

Data users reported that this lack of clarity on anonymisation processes has resulted in overly risk-averse behaviours (e.g., treating all data as personal data), and has reduced the re-use of health data and the speed of innovation due to strict definitions of anonymisation acting as a barrier to data transfers. Data users also stressed that over-anonymisation can significantly reduce the quality, usability and reliability of health data.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• Interpretation of applicable methods for anonymisation varies significantly among regional, national and European authorities, causing internal interoperability issues. • Some countries apply a stricter definition of 'sufficient anonymisation' which further limits the sharing of data for cancer research on the basis that the individual could potentially be traced and re-identified due to the rarity of their illness. • Difficulties in following the patient through the health care system when more than one care provider, each with their own interpretation of 'sufficient anonymisation', is involved in their care. • Lack of guidance on how to achieve anonymisation on broad categories of personal health data, such as medical images and longitudinal data. • Uncertainties about how and if certain types of personal health data can be accessed. Secondary impacts include delays and financial costs. • Risk-averse behaviours due to lack of clarity. Sometimes treating all data as personal data due to this lack of clarity. • Speed of innovation is reduced or impeded. • Impact on national and international projects and consortia that require exchange of health data. • Over anonymisation can reduce data quality, usability and reliability to the point that the data could potentially be inaccurate. Over-anonymisation reduces data usability in research as it is often important to do correlation studies where individual data linkage is essential. • Unclear public communication around health data use.

Policy options
Option 1: The European Commission creates legislation that includes clear guidance on the interpretation of 'anonymisation' for the European Health Data Space in collaboration with the European Data Protection Board. **Option 2:** The European Commission creates a checklist for Member States to report their anonymisation rules and interpretations. **Option 3:** The European Commission creates a common reference document that captures Member States' anonymisation practices (Option 2) and clearly communicates

countries' national level rules and interpretations. This document should be maintained and translated into all EU languages by the European Commission and regularly updated by Member States.

Option 4: Do nothing.

6.3 Barrier C: A lack of a common European interpretation of what constitutes 'pseudonymisation'

Although Article 4 of the GDPR provides a definition for pseudonymisation, data users identified a lack of guidance at national and international level on the pseudonymisation of health data leading to differing approaches to pseudonymisation, both within and across European countries. They also identified a lack of consensus on the degree of separation needed between the re-identification key and the data user for data to be considered pseudonymised, and a lack of consistency on whether pseudonymisation or anonymisation of data is most appropriate.

Data users highlighted that the use of different pseudonymisation standards and methodologies creates interoperability problems and ultimately acts as a barrier to data sharing for data users. The lack of interoperability can lead to high costs when third party involvement is required to align the data sets, or the application of additional non-standard safeguards is requested.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• Data pools may not be interoperable due to different standards/methods of pseudonymisation. • Data is sometimes required to be sent to a third party in order to align the format and combine two separate data sets. This may involve new technology, foreign to all data controllers. • Under pseudonymisation it is not possible to share some types of data for rare diseases because of semantic interoperability issues in the sector of rare diseases. • The application of safeguards adding significant cost and resource requirements. • The need to create individual solutions for each project requires more resources. • Financial costs because of repeated processes, time and resource commitments to establish all agreements and prepare the data. • In some European countries, only aggregated data can be shared for secondary use and research purposes, and not pseudonymised data.

Policy options
Option 1: The European Commission creates legislation that includes clear guidance on the interpretation of 'pseudonymisation' (Article 4) in collaboration with the European Data Protection Board. **Option 2:** The European Commission aligns assessment tools used by data protection officers (DPO), via the European Data Protection Board, with national DPOs endorsing and communicating the agreed approach. This is also suggested as a recommendation in the study published by the European Data Protection Board on the appropriate safeguards under Article 89(1) GDPR for the processing of personal data for scientific research[3]. **Option 3:** The European Commission encourages setting up a code of conduct or a certification system (e.g., Open Systems Interconnection model (OSI model)) at national level to harmonise the process of pseudonymisation. This should provide clear guidance on the level from which personal data is considered sufficiently pseudonymised (e.g., the distance from possible re-identification, how detached the third party that holds the pseudonymisation key is). **Option 4:** Do nothing.

6.4 Barrier D: A lack of a common European interpretation of what is and is not 'secondary use' of data

Data users reported that the lack of a commonly agreed definition or interpretation of what 'secondary use' of data constitutes presents a barrier to cross-border data sharing. 'Secondary use' has no basis in law in all EU Member States and there is no clear delineation between primary and secondary use.

Data users evidenced that the lack of clarity and inconsistent uses of these terms can create significant challenges when obtaining consent, where it can be difficult to interpret, what individuals have consented to. It can also lead to difficulties for ethical review boards to determine if consent has been given. This issue is also noted in the NIVEL study 'Assessment of Member States' rules on health data in light of the GDPR'[4]. Furthermore, the term 'secondary use' is inconsistently used and interchanged with the term 'further processing'. Recital 50 offer some guidance on when further processing may be compatible with the use for which the data were originally collected, however there is no conceptual clarity as to the difference between the two terms although it is clear that the secondary use described by users is often fundamentally different from further processing as described by Recital 50. Finally, some interpret the terms 'primary and secondary use of data' as data coming from primary and secondary care respectively.

[3] Kindt et al. Study on the appropriate safeguards under Article 89(1) GDPR for the processing of personal data for scientific research, Final report. EDPS/2019/02-08. (2019). Available at: https://edpb.europa.eu/system/files/2022-01/legalstudy_on_the_appropriate_safeguards_89.1.pdf

[4] European Commission, DG Health and Food Safety. 'Assessment of Member States' rules on health data in light of the GDPR'. (2021) Available at: https://ec.europa.eu/health/system/files/2021-02/ms_rules_health-data_en_0.pdf

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• Certain research studies cannot be conducted due to unclear definition of what secondary use of health data is, its purpose and whether it is compatible with what is allowed. • Reluctance from organisations to access data due to the lack of clarity on the accessibility procedures. • There is a lack of clarity on how to archive video data, and other personalised contents linked to the video data. Video data cannot be fully pseudonymised and or anonymised and hence it is unclear if and how secondary data analysis of narrative, ethnographic and video data is allowed. • Some data users reported that they treat all health data as secondary data as a preventative measure. • Data coupling to include socio-economic and behavioural population data is often extremely difficult, costly or time consuming. • Difficulties in the (re)use to clinical study data due to lack of clarity with regards to the need for reconsent from ethical review boards • If researchers have to received reconsent from patients, and such reconsent is not possible, analysis of data that could yield a therapeutic option for a critical disease or pandemic may be impossible.

Policy options
Option 1: The European Commission creates legislation that clearly outlines rules for collection, use and sharing of data for secondary use. This would include a clear legal definition of what is considered secondary use of data. **Option 2:** The European Commission and Member States agree to a high-level definition of secondary use of health data at EU level, in consultation with domain experts and the European Data Protection Board, and ensure its application. **Option 3:** Member States implement national level processes to clarify what constitutes secondary use of data domestically. **Option 4:** Do nothing.

6.5 Barrier E: European countries have national laws/rules on health and research data in addition to the GDPR

Data users highlighted that the differences in interpretation of the GDPR across countries and the existence of additional national rules can cause complications in the secondary use of health data across Member State borders. It is important to note that this statement does not refer derogations under the GDPR, but rather additional national level legislation which

applies in addition to the GDPR. This existence of overlapping acts at EU and national level has led to differences in interpretation and applications of data sharing across Europe.

Data users evidenced that these issues lead to difficulties accessing certain types of data (e.g., genomic data) and drive risk-averse behaviours. Their evidence also highlighted an over-reliance on consent including when the GDPR does not require it and when, as stated by the EDPB, it is not the most desirable method. Furthermore, different approaches to the use of consent across Europe can exacerbate these difficulties, and the need for different processing requirements across countries can hamper joint research projects.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• Conflict between the GDPR and additional national laws covering use of personal data hinders data sharing. • Risk-averse behaviour leading to reliance on consent, including when GDPR does not require it, is an obstacle for scientific research. • Difficulties to access certain types of data (e.g., genomic data or data from certain subjects) due to overly cautious and risk-averse behaviour by data controllers. • A disproportionate “stacking” of multiple overlapping safeguards required by different jurisdictions. • Where cross-border processing takes place and partners in a research consortium jointly process the data, the different national legislations have to be applied at the same time. This leads to different processing requirements, which hampers joint research projects. • Differing approaches to the use of consent worsen research difficulties to obtain consent retrospectively for secondary use of health data. • No clear, or complete lack of, rules for the use of health data for research, particularly if the researcher is not part of the healthcare provider that holds the data. This causes a broad range of issues regarding availability and conditions for the secondary use of health data. • Health data is underutilised as a resource for secondary use. This leads to inefficient use of resources, human and financial, due to the need of re-collection of the same health data for other purposes. • In countries with strong and clear legislation covering privacy aspects of research data we see that such legislation acts as an enabler for meaningful research.

Policy options
Option 1: The European Commission encourages Member States to harmonise, update or amend national legislation to remove any conflict across different acts and with the GDPR. **Option 2:** The European Commission develops a platform where ethical committees, privacy officers, data protection officers (DPO), and privacy approval bodies can interact

across Europe to share experiences and promote a more harmonised application of the GDPR.

Option 3: The European Commission develops a common reference document outlining the national laws in relation to the GDPR currently in place in Member States. This document should be maintained and translated into all languages by the European and Commission and regularly updated by the Member States.

Option 4: Workshops are organised to discuss cases and how to interpret GDPR. The workshops could be organised at EU and/or national level as differing application of the GDPR is evident at both levels. The EHDS could organise such workshops, for example.

Option 5: Do nothing.

6.6 Barrier F: European countries can set different derogations under the General Data Protection Regulation

Article 89 of the GDPR allows Member States a number of derogations from the data subject rights referred to in Articles 15, 16, 18 and 21 for scientific or historical research purposes or statistical purposes subject to appropriate safeguards where there is a basis in Member State law. Likewise, Member States can enact exemptions from the data subject rights outlined in Articles 14, 17 and 22.

Data users noted that guidance from national data protection authorities and the EDPB on how the GDPR has been implemented in different Member States, and how it should be understood, interpreted and applied in various circumstances is still forthcoming. Currently, different rules are applied in different countries, which can delay and hamper cross-border research and data sharing due to a lack of clarity about how and which overarching rules for consent and health data sharing apply.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts

- Waiting for clarification on the GDPR positions as adopted by individual Member States and for guidance from regulatory and professional organisations risks severe delays to research projects.
- Difficulties in creating an approach to the processing of personal data which would be legally compliant in all Member States. Cross-border consortia are hampered.
- Research becomes hampered as data subjects' rights can impair or make the research project impossible where no derogation was foreseen under national law.
- This challenge is made more acute with regards to the fact that the medical data at the heart of many projects are classified as "special category data" and so is subject to additional constraints as to its processing.
- Different rules in different countries create conflicts as data subjects may exercise their rights against one controller but not a joint controller in the same consortium and it is not clear what that means for overarching big data collection.

- Research projects are hampered as they prefer not to use personal data in order not to fall under GDPR, due to the delays that this would cause.
- Uncertainty over lawfulness of consent when it should also cover secondary use, especially when conducted by a separate research organisation and cross-border. Therefore, consent accepted in one country with a certain broadness, subsequent recipients and/or collection context may be valid in one country but not in another. This could create massive disruptions in data sharing.
- Broad consent in some countries does not include the export of pseudonymised information (especially from large older cohorts), limiting access and sharing of data across borders.

Policy options
Option 1: The European Commission creates legislation for the European Health Data Space which acts as the basis for certain types of processing or derogations. In this way the European Health Data Space could make use of Art. 89 but also Art. 23 of the GDPR to create harmonisation within the European Health Data Space and provide derogations important for health research.
Option 2: The European Commission ensures that the governance of the data platform is both centralised and federated within the European Health Data Space.
Option 3: The European Commission develops a reference document that captures Member States' derogations offering clear communication of national level rules and interpretation within the European Health Data Space.
Option 4: Do nothing.

6.7 Barrier G: European countries have different preferences as to the choice of legal basis for processing under the GDPR

Data users highlighted European countries' differing preferences on the choice of legal bases for data processing, as required by Articles 6 and 9, for processing personal data and processing special categories of personal data respectively, of the GDPR, which creates a barrier to cross-border collaborations and data sharing initiatives.

Data users outlined that the lack of consensus on the legal basis for processing can result in data being collected or made available under different legal bases, for example using consent in one country and using public interest or legitimate interest in another country. They evidenced that the different preferences of conditions for data-sharing among the Member States can hamper the successful implementation of trans-national research projects and pan-European initiatives. In fact, serious delays have been experienced by research institutions that have different approaches to contracting for projects and they reported the need for additional human and financial resources to maintain and monitor individual contracts and fulfil the requirements of the GDPR.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• A controller may have to apply different legal bases for the same processing where data are collected from different countries/sources. • The application of derogations from the data subject rights under the GDPR depends on the choice of the legal basis. Where exemptions to particular rights apply under one controller, they may not apply under the other, or they may only apply to part of the data. • A controller who is obliged to process under consent, but the available data was collected under a different legal basis, needs to get through derogation processes to be able to use the data with an unclear outcome. • The same is true for data types such as genomics where consent is required for processing and no derogation is foreseen in the law. • Different legal bases may apply to different datasets, but also to individual data types only. • Different Member State preferences for conditions for data sharing among Member States hampers, or could hamper, successful implementation of trans-national research projects and pan-European initiatives. • Serious delays have been experienced as each institution has a different approach to contracting for projects. • Human resources and financial costs to maintain and monitor individual contracts and to fulfil the requirements of the GDPR. • The situation is next to impossible to explain to data subjects.

Policy options
Option 1: The European Commission creates legislation with regard to the legal basis for processing specifically within the European Health Data Space. This law would apply whenever data is shared through the European Health Data Space rather than Member State law. **Option 2:** The European Commission sets up a platform for the European Data Protection Board to work with national regulators towards a united approach for the legal basis for health data sharing across all European countries. **Option 3:** The European Data Protection Board, under article 70(d) of the GDPR, creates clear guidelines (including best practice examples) on how to archive video data and other personalised content linked to the video data in compliance with the GDPR. **Option 4:** Do nothing.

6.8 Barrier H: Health data is considered sensitive data e.g., special category data under the GDPR, and is treated differently from other types of data when it comes to health data ethics, management, and use

Health data has a special category status within the GDPR (Article 9(1)) and is considered sensitive data, meaning that additional rules are applied in addition to the requirements of GDPR for processing other types of personal data, with more stringent rules for data ethics, management and use. Data users report the special category status can lead to overly risk-averse behaviours in applying the GDPR to health data, acting as a barrier to research. Data users also evidenced that it can lead researchers to resort to anonymisation of data as a mitigation measure, reducing the usability of data.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• Risk-averse behaviours, reducing the amount of potential life-saving research even in instances when legal and ethical approvals are in place. • Resorting to anonymising data as a mitigation, which reduces the usability of the data. • Excessive administrative requirements due to extra reassurances being required regarding legal and ethical frameworks. • Some countries see consent as the only possible legal basis in the case of health data, even where this contradicts European Data Protection Board guidance[5]. • The need to move certain operations outside EU countries and the associated financial and resource related costs. • Lack of clarity on the legal basis for genomic data use.

Please note that the options for this barrier have been divided into three sub-categories in order to be able to provide more specific options for practical implementation and to reflect stakeholder discussion around the need for secure environments for the processing for sensitive, special category data. The three subcategories are:

- Health data as sensitive category personal data under the GDPR
- Secure Processing Environments
- Data mobilisation capabilities

[5] EDPB. Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research. (2 February 2021). Available at: https://edpb.europa.eu/sites/default/files/files/file1/edpb_replyec_questionnaireresearch_final.pdf

Policy options
Health data as sensitive category data under GDPR
Option 1: The European Commission develops best practice guidelines including successful examples of how special category health data has been managed under the GDPR. **Option 2:** The European Commission develops a platform where ethical committees, privacy officers, data protection officers, and privacy approval bodies can interact across Europe to share experiences and come to a more harmonised application of the GDPR. Such a platform could also organise workshops on the implementation of the GDPR. Such a platform could be part of the European Health Data Space network. A good example of a similar platform is the DPO-Connect EU funded initiative that exists in Belgium. **Option 3:** All organisations dealing with sensitive health data invest in hiring or subcontracting data protection officers, beyond current GDPR requirements. These officers should also have to first complete a common European training in order to align the GDPR interpretation at this level between Member States. **Option 4:** The European Commission encourages every heath data sharing infrastructure to have a FAIR-based, open access metadata. **Option 5**: The European Commission introduces a mechanism to fast-track health data transfers due to emergency situations e.g., a harmonised GDPR derogation. **Option 6:** Do nothing.
Secure Processing Environments
Option 1: The European Commission provides clear guidance on Secure Processing Environments attributions and definitions. **Option 2**: The European Commission provides a supra-national set of Secure Processing Environments to be shared among Member States. **Option 3**: Member States sign a recognition principle between cross-border Secure Processing Environments operation and functions.
Data Mobilisation Capabilities
Option 1: Data remains within Member State boundaries for its access and analysis in national Secure Processing Environments. **Option 2**: Data can be mobilised cross-border between trusted Secure Processing Environments, during the duration of specific projects only.

6.9 Barrier I: A lack of standardised data sharing agreements for products developed by private sector providers using public health data to facilitate safe data sharing and protect public investment

Data users highlighted the lack of standardised data sharing agreements between public and private sectors for the secondary use of health data as a barrier to data-sharing collaborations and research/innovation development.

Data users evidenced that the lack of standards and disproportionate provisions leads to inefficient, time-consuming and costly tailor-made approaches and can ultimately block health data exchange. Data users also noted that they believe it is important to have equal, regulated access to health data so that fair competition principles can take place.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• If the safeguards in the data sharing agreement are disproportionate to the risk this may lead to the decision not to agree to a health data exchange. • European citizens might not benefit from health outcomes improvements originating from real-world data analysis as well as corresponding cost savings. • It is important that all data users have equal, regulated access to health data so that fair competition for the best solutions can take place. • Since there are no clear standards, to make private public collaborations possible, a tailor-made approach must be implemented to any given situation. • Loss of development opportunities based on data sharing (data economy). • Costly and long process of obtaining data for secondary use between private and public organisations. • Risk-averse approaches to data sharing. • There is no standard format for applying for data. Therefore, every organisation that has data to share does so as they see fit. • High financial impact/barrier, as the private organisations that own data request significant amounts of money for data that are being collected using public money. • Cross-border data sharing outside the EU is a barrier to conducting clinical trials in Europe for US-based pharmaceutical companies. Cloud computing services owned by a US company are legally considered unusable for storage, sharing, or analysis of European data, with major impacts on the ability of multinational pharmaceutical companies to conduct clinical trials that include European subjects.

Policy options
Option 1: The European Commission implements a framework or standard policies for data exchange for secondary use of health data from the public to the private sector and vice versa.

Option 2: The European Health Data Space terms and conditions of use indicate that research articles derived from data accessed through the European Health Data Space should be published in open-access journals.

Option 3: Member States and associated countries set national level rules for collaboration between the public sector and private industry.

Option 4: Member States and associated countries set national level rules stipulating that all data that is collected using public money, routinely or not, should be available for free after a certain time period from collection.

Option 5: Public and private sectors inform citizens on how and where their data was used.

Option 6: Do nothing.

6.10 Barrier J: The use of different interoperability standards across Europe makes comparisons and sharing data and research results challenging

Data users reported a number of issues regarding semantic interoperability. The use of different terminologies has caused delayed responses to urgent requests, has caused difficulties in data linkage and has often led to misunderstandings and diagnostic errors, for instance in the case of rare diseases.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts

- Difficulty in analysing the impact of health initiatives due to differences in the interpretation as well as differences in medical practices and specialist structures.
- The current SNOMED CT-Orphacode map does not capture all of the individual rare diseases (85% coverage) and has no provision for coding unknown rare diseases or for flagging a disease as being rare, meaning not all rare diseases can be counted, and not all rare diseases can be aggregated for analysis by rare disease groups.
- Different terminologies hinder speedy responses to urgent needs, and lead to inefficiencies when multiple parties do the same thing.
- Common models do not work well with complex data that are not simply observational.
- Semantic mapping to international ontologies and terminologies impacts on the data being able to be used from different data sources.
- Missing support of relevant ontologies and taxonomies, such as SNOMED CT, and relevant standards for Health Information Exchange, such as HL7-FHIR, by medical practice software, hospital information systems and population health information systems.

Policy options
Option 1: The European Commission adopts a single interoperability standard for the European Health Data Space. **Option 2**: The European Commission recommends adoption of a predetermined list of a few common interoperability standards including terminologies, ontologies and classification systems. **Option 3**: The European Commission develops a reference document, maintained by the European Health Data Space central service structure, outlining the standards used by each Member State to describe health data. This would include clear documentation of the ontology/taxonomy used so that it would be possible to create common data models or ways of translating one into the other. **Option 4**: The European Commission promotes the use of ontology servers, open published tools and tool sharing (EU and national level) with international mapping. **Option 5**: Member States create a national infrastructure and use a common standardisation protocol. **Option 6**: Do nothing.

6.11 Barrier K: Poor data management procedures reduce the ability to reuse data

Data users report that poor health data management is a major barrier to health data exchange for secondary use between institutes within the same country, and across European countries. The lack of adherence to the FAIR principles is reported to be one of the main sources of the issue.

The impacts of this barrier as experienced by European data users are summarised in the grid below, followed by the options to address it, based on the best practices developed with and informed by data users.

Impacts
• Poor data management causes loss of value of information being generated because of missing or inconsistent data entry. • Inconsistent data access procedures cause difficulty ascertaining where to access the data and slow or delayed data access. • Poor management makes re-use of data time demanding. • Lack of transparent health research project approval process. • Slow processes for data integration for large projects. There is no developed system for federated data integration for public data, hospital data (EHR), genomic data. • Poor data management can impact international benchmarking if it is not clear the data sets yield comparable information (clear metadata), causing time and financial costs. • Academic institutes often argue that data collected pre-GDPR can no longer be used or shared for retrospective studies, as they may be concerned whether consent was collected or explained in accordance with the GDPR at the time of the original data

collection. It is not always clear whether this hesitance is due to inefficient data management now and in the past, or whether this is due to a GDPR interpretation issue. • GDPR and the mind-set that it has encouraged has made it difficult to obtain individual-level data. This has led to some poor research on important topics especially those to do with inequalities between selected groups.

Please note that the options for this barrier have been divided into two sub-categories in order to be able to provide more specific options for practical implementation and to reflect stakeholder discussions on poor data management. The two subcategories are:

- Poor data discoverability and access
- Lack of common data quality assessment framework

Policy options
Poor data discoverability and access
Option 1: Every national data hub and every data infrastructure must have public metadata catalogues, using a standardised template outlining minimum standards to ensure findability. **Option 2**: The European Health Data Space central service structure creates a European level list of existing health data sources and data controllers, with a research focus and with good descriptive metadata for each data source. **Option 3**: The European Commission recommends the use of standardised terminologies, common data models and coding systems to ensure interoperability. **Option 4**: The European Commission creates a standardised data dictionary with definitions and terminologies which data controllers must abide by. **Option 5**: Every national data hub and every data infrastructure uses common data models for structural metadata templates.
Common data quality assessment framework
Option 1: The European Commission establishes a European Health Data Space data quality assurance governance structure, which could ensure adherence to good data management procedures across Europe and provide professional guidance on health data quality assurance. A national data permit authority could audit the data holder institutions and their data sets in accordance with an EU data protection authority or another EU data quality assurance body. **Option 2:** The European Commission develops a common framework as a requirement for a data controller institution to be deemed a trustworthy party in the European Health Data Space. An EU body, in collaboration with the national data protection authorities, would promote the implementation of such a framework, incentivise the adoption of measures for improvement and supervise the effective adoption and its maintenance. **Option 3: The** European Commission sets an obligation for data holders to conduct a regular data quality audit at institution level. An EU body along with the national data

protection authorities could establish the procedures and entitle third parties to perform those audits.

Option 4: An EU body along with the national data protection authorities sets up a benchmarking and promotion mechanism to incentivise data holders to upgrade their data collections and procedures and to make them available to the European Health Data Space.

Option 5: Member States develop consistent policies which define access protocols at national level and/or per institute.

Option 6: Member States and associated countries provide the possibility to have secure environments which allow researchers to access and/or analyse the data.

Option 7: All Member States and associated countries establish a National Research Ethics Committee.

Option 8: The European Commission develops an identification system at EU level, similar to the EU research funded projects Participant Identification Code, which could help to streamline data access requests across Europe for trusted organisations.

Option 9: The European Commission encourages Member States and associated countries to invest in user-friendly interface for EHR-systems, which would support and facilitate the registration of (structured) data according to the workflow and facilitate the transition to a 'circular-health-data'.

Option 10: The European Commission develops additional guidance on data management and training on epistemology and technology of research (capacity building at EU level) and to invest more in robust IT infrastructures (national). For example, implementation of a Data Management Certificate.

Option 11: Do nothing.

7 Discussion

Literature review

The initial literature review highlighted that European data users experience a wide range of barriers to cross-border data sharing and enablers to overcome them. The consolidation of individual reports of barriers to data sharing, with a specific focus on data users for the first time, facilitated a new and deeper level of analysis on the impacts and needs of this stakeholder group.

Of the 89 barriers identified in the literature, 33 were classified as barriers relating to data management and 18 as legal barriers. Together these two thematic areas represented 57% of all the barriers identified. The remaining 43% of barriers identified by data users related to infrastructure, trust and transparency, resource issues and ethical aspects.

Furthermore, many of the 89 barriers identified highlighted the same or similar issues. For example, 29 of the barriers stemmed from a lack of sematic interoperability and differing interpretations of key terms.

The 79 enablers broadly corresponded to the same thematic areas but ranged greatly on the level of intervention suggested by data users. A key area of tension was whether legislative

or non-legislative options would be most effective and whether the European Health Data Space should aim for full harmonisation or alignment of Member States and associated countries' data sharing interpretations and approaches within the Data Space. These tensions are reflected in the options which range from high to low invention intensity.

Overall, the initial findings suggested that data users were most regularly affected by legal and data management barriers to data sharing caused by misalignment of interpretations and implementation.

Framework

In order to test this hypothesis and explore the impact further, a data user framework was developed. The framework provided an overview of the thematic areas within which data users experience barriers to cross-border data sharing. A further prioritisation exercise to understand which barriers had the most significant impact on data users, resulted in six of the final eleven priority barriers being legal issues. This prioritisation exercise also allowed us to target our work to identify best practices to address the most problematic barriers for data users. The final barrier list provides an overview of the most significant barriers to data sharing by both issue and theme.

Case studies

Based on the framework, extensive stakeholder engagement was carried out to collect over 100 case studies to provide real-life examples and evidence of the impact of these barriers on data users, as well as to gather best practices and stakeholder recommendations to address the barriers.

Of all the case studies, 52% were submitted regarding legal barriers, 30% regarding barriers caused by data management, 13% technical issues and 5% linked to barriers to trust and transparency. The comparative level of case study submissions by data users further supported the finding that the legal barriers identified to data sharing are especially challenging for data users.

It is important to note that the sample size of responses per barrier is not large enough to make conclusions between specific barriers that are more important for different user categories (i.e., researchers or policymakers), as outlined above. The results from the different data user categories were often similar. Therefore, the analysis combined the perspective together as data users, whilst acknowledging the nuances in their perspectives.

Impacts

Data users were encouraged to share the impacts of each barrier discussed within their case study to facilitate a better understanding of the specific issues as well as potential best practices and suggestions to support the development of policy options. An analysis of their responses shows a pattern of overarching impacts across all four thematic areas. This includes different applications and interpretation of legislation and data management processes causing interoperability, access and analysis issues and ultimately having an adverse impact on patient care. Data users also report that a lack of clarity and harmonisation has created a risk-averse culture and reduced the speed of innovation unnecessarily, to the detriment of European patients. In summary, the effects of barriers to data sharing experienced by data users result in health data being underutilised for secondary use and

reduces the benefits for all. Data users suggested a wide range of best practices and suggestions to address these barriers, ranging from low to high intensity interventions. Their input, and the intensity range are reflected in the policy options put forward in this report.

Policy options

Based on the best practices from the case studies, policy options to address the identified barriers were developed. It is not within the scope of this report to suggest one preferred solution for each barrier, rather to present a set of options, developed with and by data users, for wider TEHDAS tasks and the European Commission to consider. Following stakeholder input, all options have been refined and tested through consultation, expert interviews and policy development. These options are intended to be practical and actionable and following publication of this report they will be taken forward as set out in the section on 'Next Steps.'

8 Next steps

The aim of this task was to provide the evidence base and data user perspective to inform wider TEHDAS tasks and work packages as well as the development of the EHDS infrastructure, governance and legislation, including:

Work package 5, Task 5.2 will use the case studies and policy options to develop guidelines/recommendations for European countries to consider when planning national legislation to enable cross-border exchange and secondary use of health data (deliverable 5.2).

The case studies and policy options will also provide the evidence base for wider **Work Package 5** deliverables and corresponding milestones, including developing recommendations for best practices for EU cross-border exchange including data access and data permit processes in different national settings (deliverable 5.2) and options for governance models for the EHDS (deliverable 5.4).

Work Package 4: Task 4.1 will initiate country visits to map the data management processes and organisation in relation to the future EHDS. Task 5.1 outputs will provide information on the health data management processes in some countries. The literature review will also be used to identify key reports for the country visits.

Work Package 6: The case studies and policy options on semantic interoperability and data quality will provide the evidence base for work package 6 work to develop the EHDS data quality assurance framework for secondary use of real-world health data and the EHDS' semantic interoperability framework (deliverable 6.2).

Work Package 7: The case studies and policy options will support WP7 work on assessing user's expectations (milestone 7.2) and the development of options for the services and services architecture and infrastructure for secondary use of data in the EHDS (deliverable 7.2).

Work Package 8: The case studies will provide information on the wider landscape in order to situate the iCitizen work (deliverable 8.1) and a number of the policy options suggested by data users may also be applicable to WP8's work to identify and develop methods to support data altruism in the implementation of national health data spaces (deliverable 8.2).

9 Conclusion

In conclusion, this report presents the perspectives of data users (researchers and policymakers) on the secondary use of health data within the European Health Data Space (EHDS). It consolidates the results of the literature review, data sharing framework, stakeholder case studies and expert interviews. The list of impacts and policy options to overcome the priority barriers identified by data users provides the evidence base to inform wider TEHDAS tasks and work packages as well as the development of the EHDS infrastructure, governance and legislation. Barriers to data sharing caused by semantic and legal interoperability were identified as having a significant impact on data users and will need to be addressed as a priority within the EHDS to ensure its success.

10 Annexes

Annex 1 – Literature review key terms and selection criteria

1.1 Setting eligibility criteria

A set of inclusion and exclusion criteria were selected, and a search methodology was established. Inclusion criteria:

- Published in the last 5 years (i.e., after 2016),
- Published in English,
- Study conducted in the EU and/or associated countries, and
- Limited to documents, articles, reviews, systematic reviews and meta-analysis.

Excluded studies were those that were from outside the EU or associated countries, not in English, conference papers, book chapters, conference reviews, short survey, notes, editorials, books, clinical trials and randomised controlled trial (RCT).

1.2 Search criteria and key terms

Table 1 presents the search terms used to identify relevant studies in PubMed and Scopus databases. The search yielded 39 and 90 outputs, respectively. When duplicates were removed 121 papers remained.

Table 1. Final search terms for Pubmed and Scopus data bases.

Pubmed:	**Scopus:**
"barriers" OR "enablers" OR "SWOT" OR "Cross border" OR “Cross-border” OR “Gaps” OR “best practice” OR “weakness” OR “strength” OR “opportunities” OR “threats” AND "sharing" OR "share*" OR "link*" OR " access" OR “transfer" AND "Health data" OR " Health Information" AND	(("barriers" OR "enablers" OR "SWOT" OR "Cross border" OR "Cross-border" OR "Gaps" OR " best practice" OR "weakness" OR "strength" OR "opportunities" OR "threats") AND ("sharing" OR "share*" OR "link*" OR " access" OR "transfer") AND ("Health data" OR " Health Information") AND ("secondary use" OR "reuse" OR "re-use" OR "secondary purpose" OR "GDPR" OR "data protection" OR " governance" OR " policy" OR " guideline"))

"Secondary use" OR "reuse" OR “re-use” OR "secondary purpose" OR "GDPR" OR “data protection" OR " governance" OR " policy" OR " guideline"	

Annex 2 - Governance and health data management systems in European countries

Austria

a. Overview

The Austrian Ministry of Health is responsible for oversight and aggregated data for policymakers and health authorities, while maintaining and developing federal data systems for special purposes, such as the epidemic reporting system, or drug addiction treatments. Databases developed by the Ministry further allow for processing and analysis based on aggregated data, with the Ministry acting as data controller. Data originates from a range of healthcare providers.

Other bodies, such as ELGA – a legal entity owned jointly by the federal administration, federal states, and the social insurance developing the Austrian electronic health record system – is tasked with handling, exchanging, and making accessible electronic health records. These records are legally not available for research, and there is no cross-border exchange.

In addition, several other entities and organisations hold personal data relating to citizens' health. The Austrian social insurance processes data related to health insurance cases. The Institute for Public Health and Food Safety (AGES), collects, maintains, and processes data related to food and medicine safety. The National Institute for Health Research (GÖG) provides pandemic data to verified medical universities after a standardised application process.

b. Accessing health data in Austria

Access to health data must have a legal basis. If such a basis is provided, standardised processes allow access to such data to research institutes and universities. The health data which are shared are typically pseudonymised. There are a variety of access mechanisms and monitoring processes, dependent on each organisation's own governance processes. For example, requests to access pandemic data held by GÖG overseen by a committee with delegated authority.

Belgium

a. Overview

In Belgium there are a number of actors within the health data system. Key organisations include:

- **Intermutualistic Agency**: health care prescription data
- **Statbel:** mortality and cause of death data
- **Belgian Cancer Registry**: cancer diagnosis

- **FPS Health**: hospital discharge data
- **Intego (KU Leuven)**: primary care (GP)
- **Sciensano**: health surveys, surveillance, and rare disease registries

The federal health data system is eHealth, a platform that includes the following organisations:

eHealth Platform Organisations	
National Institute for Sickness and Invalidity Insurance (INAMI)	Federal Public Service Public Health, Food Chain Safety and Environment
E-Health platform	Scientific Institute of Public Health
Brussels Health Network – Abrumet asbl	Federal Agency for Medicines and Health Products
Medex	Data protection authority
Agentschap Zorg & Gezondheid	Federal Center of Expertise for Health Care (KCE)
Walloon Health Network (RSW)	Agency for a Quality Life
Zorgplatform collaborator	Vlaams Ziekenhuisnetwerk KU Leuven
National Intermutualist College (CIN)	Mutual funds
The League of Users of Health Services (LUSS)	ZNA – care portal

These organisations can act as controller, processor or neither depending on the specifics of the data processing. There is no "fixed" definition of who can act as data processor or controller because this depends on specific applications. Data comes from citizens (directly or indirectly e.g., from registers, surveillances, other databases). Organisations can have different types of data and most have several types including health records, survey data, biological material, and samples.

b. Accessing health data in Belgium

Access to health data is granted through collaboration agreements between the data holder and the organisation requesting access. In some circumstances it is necessary to submit an application to the information security committee.

Denmark

a. Overview

The two main national bodies that host health data are Statistics Denmark, which stores data about the wider Danish population, and the Danish Health Data Authority (Sundhedsdatastyrelsen), which hosts disease registers and data bases with health-related information. Statistics Denmark is a public independent agency and holds copies of register data and can extract health data and combine it with social conditions when the researcher requests it.

All data is exchanged via the platform Sundheddatanettet. Data is not stored there but it is a secure space where you need authentication and approval to be linked up through VPN-access so that you can exchange data. MedCom is responsible for developing and setting standards for data exchange and testing supplier products before they are released to ensure data compatibility. Sundhed.dk's two-year strategy intends to open up safe spaces for storage of citizen generated data, which can potentially be marked as available for research too, but this is not operating yet.

b. Accessing health data in Denmark

Researchers can apply for access to data locally with data custodians, or for the whole country through the Researcher Service (Forskerservice) at Serum Institute (when it is health data only) and through Statistics Denmark, if the researcher wants to combine health data with other data types. The Danish Health Data Authority holds all health registers and provides research support service (Forskerservice) for researchers who wish to access health data. It is also responsible for national coordination of data exchange systems and infrastructures for the provision of healthcare. The Danish Clinical Quality Program (RKKP) is the cross-regional network organisation of the five Danish regions that constitutes the infrastructure of clinical quality registries and coordinates access to the data for researchers.

Decisions regarding access are made by the steering group of the individual database. There is a fee for accessing data for research that must be paid to Statistics Denmark, the Serum Institute, or DAK-E. The fee covers the hours spent on setting up the specific data set, and for DAK-E it also covers the commercial vendor fee. It does not cover the cost of the infrastructure. Registry data is available for research with no informed consent ("solidarity by law").

In Denmark there is a differentiation between clinical access points and research access points. Sundhed.dk is the access point to electronic health records for patients and for health professionals for clinical purposes. A researcher needing data for research has several access points and can go to the Danish Clinical Quality Program (RKKP) for quality databases, the Serum Institute for health data, and to Statistics Denmark for registry data combined across sectors.

Primary care data must be accessed through the municipalities (for homecare and nursing homes) and DAK-E/KIAP from the Danish Quality Unit for General Practice for GP-data. Sundhed.dk is an independent agency governed by the Regions and the Government and contains the national electronic health records. At the sundhed.dk platform patients can access personal health information from electronic health records, laboratories, personal

choices (e.g., organ donor), and the national patient registry. The patients can access their record, but they cannot report data or control the data. Health professionals also have access to the electronic health records.

The National Biobank, hosted by the Staten Serum Institute, and the Regional Biobank Program provide access to tissue samples. The National Genome Centre provides access to genomic data. The Health Act specifies that all genomic data from comprehensive genetic analyses is stored in a national genomic database and that patients have the right to opt-out of further use of the data.

Statistics Denmark has been involved in several working groups to facilitate data exchange between different countries. Data from Statistics Denmark is as a main rule only available for Danish researchers, but foreign researchers can get access to micro data through an affiliation to a Danish authorised environment. The Danish Health Data Authority applies the same rules.

There is a fee for accessing data for research that one must pay to Statistics Denmark, the Danish Health Data Authority, the Serum Institute, or DAK-E (for GP data) but that only covers the hours spent on setting up the specific data set, and for DAK-E also the commercial vendor fee. It is not the cost of the infrastructure.

The capital and Zealand region of Denmark use the EPIC systems. The health data is structured in an SQL database where data is stored in both EPIC defined keys and locally defined keys. External access to health data requires an approved research project.

Estonia

a. Overview

In Estonia, the Health Information System (HIS) database contains records relating to health care, including contracts for the provision of health services, health statistics and for the management of health care. The database was established by the Health Services Organisation Act. HIS enables the exchange of information between doctors by connecting IT systems for health services. The HIS gives doctors access to a selection of a patient's health information and provides timely, critical information to ambulance services. The data controller of the Health Information System is the Ministry of Social Affairs.

Health care providers are required to submit the following data:

- Waiting lists
- Medical images
- Health services provided to patients
- Management of health care, including for maintaining registers concerning the state of health established based on law.

The composition of the data, such as documents, conditions, and procedure for the preservation of the documents to be forwarded to the HIS, are established by the Ministry of Health and Labour.

b. Accessing health data in Estonia

Patients can access their personal data held on HIS. In order to protect a patient's life or health, a health care provider may delay forwarding data to the HIS to allow patients an opportunity to examine their personal data with a health care professional.

Health care providers and third parties involved in the provision of health services have access to the personal data in HIS for entry into and performance of a contract for the provision of a health service.

The basis for collecting data is context dependent:

- Within the health system patients must opt out. Patients' data is collected by default for all healthcare services and there is an assumption that a person agrees to the processing of their health data when using the services.
- Outside the health system individuals must opt in. This means data subjects' consent is necessary for the processing.
- Patients have the right to prohibit the access of a health care provider to their personal data in the HIS. The health service provider must be registered in the National Registry of activity licenses for provision of health services and the National registry of health care professionals.
- When researchers wish to access data for scientific research, they must apply for access to the personal data held on HIS to the controller of the HIS. The ethics committee of the HIS assesses whether the release of personal data from the HIS for the purposes of scientific research or statistics is justified. The assessment of the ethics committee is not legally binding for the controller, and the controller grants authorisation.

Finland

a. Overview

The data comes from local and regional health service providers, they are registered owners or data controllers (mainly municipalities responsible for primary health care or hospital districts, university hospitals responsible for special health care or from national registers/data controllers) like the Finnish Institute for Health and Welfare, Social Insurance Fund, Finnish Centre for Pension, Statistics Finland. Data comes digitally from electronic health record (EHR) systems to national systems and registries. The National Electronic Health Records archive system is available for secondary use of health and social data.

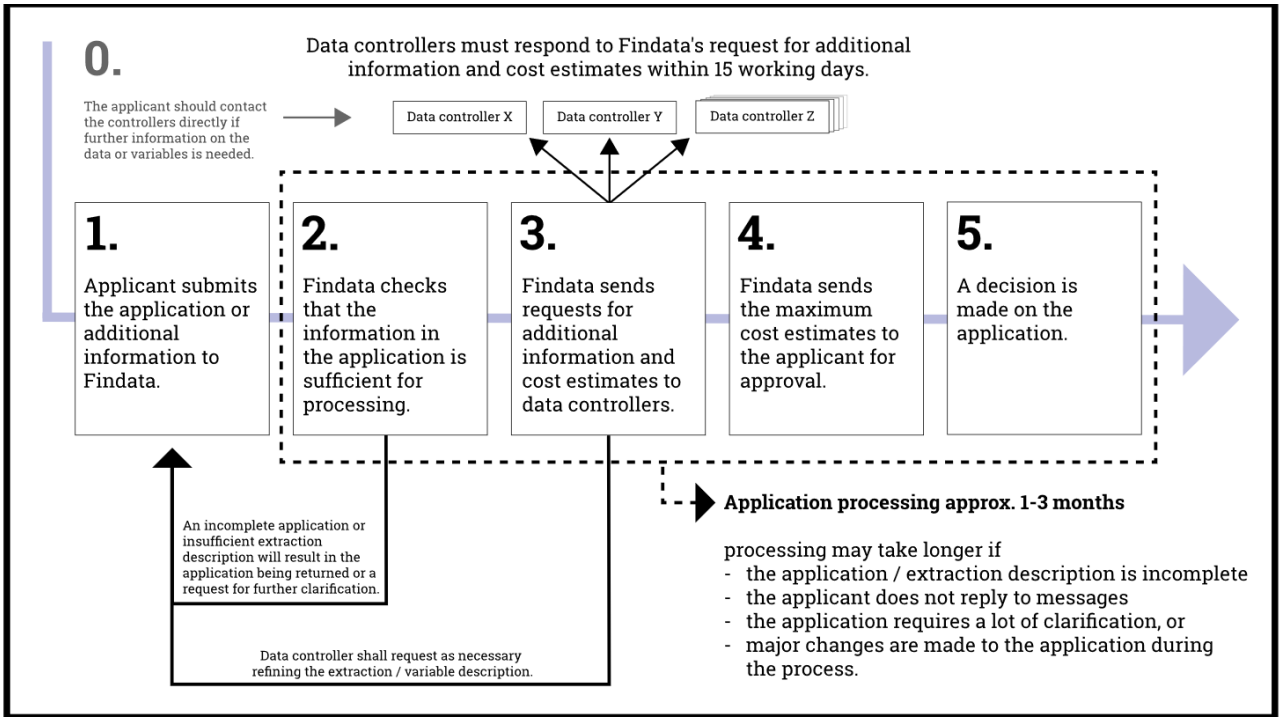
Data from 11 Finnish Biobanks is available, alongside socio-economic data from income registries, family and household data from a population registry. The organisation has social and health data: client and patient data EHRs, prescriptions, referrals, medication data,

patient flows, social care client data and social insurance data, income data, household data, family structure data, and other economic data and statistical information.

b. Accessing health data in Finland

If you need health data from only one data controller or register keeper, you need to ask for permission from that data controller or registry. If you need data from several and different sources and data controller on the national, regional, or local level, you need to ask for permission from FinData, a national social and health care data permit authority. FinData both gives the permit and collects, combines the data from the different data controllers and transfers the combined needed dataset into safe and secure ICT environment to customer who can have access to anonymous or pseudonymous data. Aggregated data can be delivered to user's own ICT environment.

Figure 1: The monitoring and data access process in Finland



Greece

a. Overview

EOPYY, the National Organisation for the Provision of Health Services, holds electronic health record data for patients through the e-prescription IT platform. Since 2011, Greece holds an e-prescription Information Platform for medicines, medical tests and materials needed to be prescribed. The e-prescription procedures are obligatory for health providers (doctors, health organisations, pharmacies) either in the public or private Health Sector, under public insurance laws. These health data do not include diagnosis and test results but only the prescribed medicines, tests, and materials.

EOPYY also holds data for all health providers with which they have legally binding contracts. Health Providers of the public sector must belong to EOPYY. Health Providers of the Private Sector cooperate with EPOYY through legally binding contracts.

HDIKA AE: is the public organisation which is the owner of the e-prescription IT Platform.

Hospital Information System: each hospital (either public or private) has its own information system, so each hospital produces, holds, and controls data of all types. These Health Organisations share regulated data with EOPYY (National Organisation for the Provision of Health Services) through e-prescription and other IT platforms. The data does not include diagnosis or test results. These organisations may exchange data under the GDPR and national laws grid for health data with other organisations, mainly for research and/or governance purposes.

Primary Health Care Units (either in Public or in the Private sector): Produce and hold health data. They share regulated data with EOPYY through the e-prescription IT platform.

b. Accessing health data in Greece

For a citizen, in order to obtain health data from a health organisation (e.g., a hospital) a citizen should lawfully apply for a copy of their health file. This procedure refers only to a person's own health data and does not include electronic data. The e-prescription IT platform is not yet open to citizens but only to health providers.

The health organisation which produces and holds the data is the only responsible entity for the data. Only under the strict national law grid for data protection and the GDPR rules and after an application of interest and purpose can someone (person, organisation, etc.) gain access to data.

Ireland

a. Overview

The Health Information and Quality Authority (HIQA) has a statutory remit to develop standards, evaluate information and make recommendations about deficiencies in health information under the Health Act 2007. The HIQA oversees the following data collections:

- National Screening Service - BreastCheck
- Hospital Pricing Office (HPO) – Hospital In-patient Enquiry (HIPE) scheme
- Health Service Executive (HSE) – Primary Care Reimbursement Service (PCRS)
- Health Protection Surveillance Centre (HPSC) – Computerised Infectious Disease Reporting (CIDR)
- HSE – National Incident Management System (NIMS)

HPSC disseminates information and data from CIDR through a wide variety of methods to ensure that infectious disease data and information is accessible to a wide range of

stakeholders. For example, weekly, monthly, quarterly, and annual reports are published online on the HPSC website. Other outputs include frequent social media posts, presentations, and scientific paper publications. HPSC have a number of national KPIs in relation to the dissemination of CIDR data which ensures the timely publication of key infectious disease data on a weekly basis. At a local level, CIDR data is used to manage infectious diseases. Within hospitals, medical staff and management can use surveillance data for audit and research purposes.

At a national level, the data is used to trend incidence and burden of infectious disease regionally and nationally, as well for planning services. In addition, CIDR data is used to enable Ireland to meet its obligations in reporting notifiable infectious disease data to international agencies such as the European Centre for Disease Control (ECDC), the European Food Safety Authority (EFSA) and the World Health Organisation (WHO). For example, the data is submitted to ECDC through the European Surveillance System (TESSy) and is used to analyse and disseminate surveillance data on infectious diseases in Europe.

b. Accessing health data in Ireland

On the HPSC website, the publications page is dedicated to disseminating a variety of reports which provide website visitors with access to summary statistical data on the range of disease topic areas monitored on CIDR. The use of information is monitored by an Information Officer and the use of data is reported in a monthly and annual report on the impact assessment of outputs. There is a process in place for assessing and processing external data requests at a national level. The protection and disclosure of CIDR data is subject to the legal remit of the Health Act 2007 and data protection legislation. The CIDR National Peer Review Group reviews requests for data from CIDR and the purpose for which it is requested. This purpose needs to be in line with the reason that the information was originally collected, that is, the surveillance, management, prevention and control of the notifiable infectious diseases and their causative organisms. To ensure that this information is protected and only disclosed appropriately, application to the CIDR National Peer Review Group is required for CIDR data requests from third parties and from CIDR partners seeking access to CIDR data beyond their current access level. The CIDR National Peer Review Group provides a clear procedure regarding the application and assessment process for accessing and using CIDR data. However, information relating to this group, or the formal data request procedure is not available online on the HPSC website.

Moldova

a. Overview

The National Bureau of Statistics (NBS) is the central administrative authority which manages and coordinates the activity in the field of statistics from the country. The NBS works independently or in collaboration with other central administrative bodies to approve the methodologies of statistical and calculation surveys of statistical indicators. The NBS ensures these methodologies are in accordance with international standards, especially those of the European Union, and with the advanced practice of other countries, as well as considering the peculiarities of the socio-economic conditions of the Republic of Moldova. In addition, the

NBS organises programmes of statistical works, annually approved by the Government, statistical surveys regarding the situation and economic, social, demographic development of the country and collects, processes, stores and disseminates statistical data.

A component part of NBS is the General Division for Social and Demography Statistics with Social Services Statistics Division, which produces statistical indicators and provides statistical data and information on various social issues, such as health, justice, public utilities, social protection and assistance, gender statistics, etc. The Social Services Statistics Division processes and controls data coming from the regional institutions working in the field of social services statistics. Statbank “Health protection” has health records, collected in electronic format through the "e-Reporting" portal.

The Ministry of Health, Labour and Social Protection provides data to NBS and has an Internal Audit Service, which conducts audit activities in subordinated institutions. The National Agency for Public Health (NAPH) is an administrative authority subordinated to the Ministry of Health, Labour and Social Protection, and is responsible for maintaining and managing the national database of health statistics. Its basic functions include the collection, standardisation, and analysis of statistical information on public health received from territorial subdivisions including public health centres, and the creation of automated systems for the collection of operational information on the population’s health.

Many other separate information flows reflecting activities within different national health programmes, and in state surveillance of public health, are managed by the NAPH: Transplant, Tuberculosis, AIDS, etc.

The Family Doctor's Centre coordinates the activity of the primary medical assistance in the territory, performs the centralisation of the statistical medical data and submits reports of their activity directly to the public health centre. The National Health Insurance Company (NHIC) is a state non-profit-making body with financial autonomy and manages a separate information system for monitoring of individuals covered under medical health insurance, oversight of contributions and economic aspects of health service provision. The NHIC covers the whole territory of the Republic of Moldova through territorial agencies, coordinating and supervising their activity within the existing legal framework.

b. Accessing health data in Moldova

Moldovan citizens have free access to the health data within the Statbank “Health protection” on the web page www.statistica.gov.md. The access to data in our health data system within the National Bureau of Statistics is free (www.statistica.gov.md). The web page www.statistica.gov.md permits to download free “statistical yearbooks of the health system”.

Sweden

a. Overview

The National Board of Health and Welfare is the data controller for health data registers in the field of health care and social services and the cause of death register. The registers form

the basis for the official statistics in the field of health and diseases, health care, social services and causes of death.

Swedish National Quality Registries is the data controller and data processor for registries containing individualised data about medical interventions, procedures, and outcomes. They are integrated into clinical workflows and have the capacity to generate data in real time. Each registry is supported by an organisation of health care professionals and patient representatives. They are jointly responsible for developing the registry.

Swedish eHealth Agency is the data controller and data processor for several registries and databases that link healthcare, pharmacies, and patients. The eHealth Agency facilitates the work of healthcare and create the conditions for better health.

Sweden's healthcare system consists of 21 regional healthcare authorities and different healthcare providers and the use of different journal systems with no or very few interconnections in between. The journals act as data controller. Data is supplied by health care practitioners and consists of health records for in-patient care, diagnoses, and pharmaceuticals.

Registerforskning.se acts as a data controller and is operated by the Swedish Research Council to provide researchers with information on existing registers, as well as support during the process of register-based research. Information on each part of the process of identifying, requesting, and using register data. A metadata tool called RUT (Register Utiliser Tool) enables efficient searching and matching of metadata in registers is also available.

b. Accessing health data in Sweden

Data access is primarily governed by the Patient Data Act, the General Data Protection Regulation (GDPR) and the Public Access to Information and Confidentiality Act which lay out how personal data may be used, and medical records are to be handled. The healthcare provider must have its own procedures that complement how and when the data may be used.

The National Board of Health and Welfare is the data controller for health data registers in the field of health care and social services and the cause of death register. The registers form the basis for the official statistics in the field of health and diseases, health care, social services and causes of death.

The data in the National Board of Health and Welfare's health data register and registers in the field of social services are covered by absolute confidentiality. From the registers, personal data can only be disclosed for research and statistical purposes and data that cannot be directly attributed to the individual. In these cases, the information may be disclosed if it can be disclosed without injury or harm to the person concerned or to any related party.

The registers are protected by confidentiality, but data may be disclosed after special examination which includes an application to the Swedish Ethical Review Authority. Each disclosure requires a formal written decision based on a special confidentiality review in which the National Board of Health and Welfare investigates whether there is legal support for breaching confidentiality under the Public Access to Information and Confidentiality Act

(24:8). This applies to new cases as well as when ordering updates to ongoing projects where the National Board of Health and Welfare has previously disclosed data.

In Sweden access depends on the purpose. For clinical use, medical staff have access to the information in the NPÖ, provided the individual gives his or her consent. All research where individuals are directly or indirectly involved or affected requires the approval of an agency (Etikprövningsmyndigheten) tasked with reviewing the ethics of the proposed research.

Health care, as well as health data, is managed separately by 21 different regions and to some extent, mostly for home care for the elderly, by 290 municipalities. Some of the health care for which the regions are responsible is outsourced to private operators who, in some cases, use the same electronic medical records (EMR) and other systems as the region-operated health facilities. In other cases, they use their own systems. In all these cases, they are controllers.

The National Patient Overview (NPÖ) allows clinicians to access (=view) limited, predefined sets of information in the EMR-systems in other regions. Separate from the EMR-systems and all other care and care management systems in the regions, there are some 100 freestanding, mostly diagnosis-related national "quality registries" in which interventions and outcomes in the different medical domains are manually recorded. The latter are used for monitoring and research within, but because of stovepipe nature not across, medical domains. The National Board of Health and Welfare (Socialstyrelsen) keeps six national registries, on cancer; patients treated in hospitals and specialised care facilities; births (including medical data); drugs sold over the counter; dental care; and care carried out at county level.

United Kingdom

a. Overview

The structure of UK health data systems is complex with multiple organisations collecting, holding, and sharing data across devolved nations. Individual NHS hospital trusts, and GPs (primary care) act as data controllers for the patients they treat. There is also a small number of private and voluntary sector providers of healthcare, sometimes commissioned or funded to provide NHS services. Local authorities and private providers of social care also hold data on recipients of social care in residential homes or domiciliary settings. Academic institutions and pharmaceutical companies running clinical trials and cohort studies also hold bespoke data about participants in their research. Health is a devolved matter in the UK, so each of the home nations has slightly different arrangements for managing health and care information, particularly in relation to secondary purposes. In England, a number of the Department for Health and Social Care's Arms Lengths Bodies also hold health data. In particular:

NHS Digital is the national information and technology partner to the health and care system. NHS Digital has responsibility for standardising, collecting, and publishing data and information from across the health and social care system in England.

NHS England also holds a range of health data – to support a range of secondary health care purposes such as service planning and population health management; it also established a COVID-19 Datastore to support the pandemic response.

Public Health England has powers to collect patient data in relation to communicable disease surveillance and other risks to public health (e.g., to support the administration of immunisation programmes). It is also responsible for the National Cancer Register.

Furthermore, UK Biobank is a large-scale biomedical database and research resource, containing in-depth genetic and health information from half a million UK participants collected with their explicit consent.

Genomics England (GE) was set up to deliver the 100,000 Genomes Project and clinical, laboratory and health data flows from a number of NHS, Social Care, and research organisations to GE.

Clinical Practice Research Datalink (CPRD) is a real-world research service supporting retrospective and prospective public health and clinical studies. CPRD collects anonymised patient data from a network of GP practices across the UK. Primary care data is linked to a range of other health related data to provide a longitudinal, representative UK population health dataset.

Health Data Research UK, the national institute for health data science, runs the Health Data Innovation Gateway. This portal provides a common entry point to discover and request access to UK health datasets.

The data collection and sharing landscape varies between devolved administrations, with some organisations (e.g., UK Biobank) operating across the four nations, whereas others (e.g., NHS Digital) are England only.

Processes for accessing health data vary according to organisation. However, all operate within the UK legal framework. All use of personal data in the UK is subject to the following data protection legislation:

- UK General Data Protection Regulation (UK GDPR)
- Data Protection Act 2018 (DPA)

The UK GDPR establishes the basis for sharing personal data (that is data which directly or indirectly identifies a living person). The DPA puts those safeguards into UK law. The legislation provides several key protections and safeguards for the use of an individual's data as set out below.

b. Principles for sharing data

Sharing of personal data in the UK has to follow strict rules and must follow the seven key data protection principles set out in the UK GDPR. These provide that personal data must be:

- used fairly, lawfully, and transparently.
- used for specified, explicit purposes.

- used in a way that is adequate, relevant, and limited to only what is necessary.
- accurate and, where necessary, kept up to date.
- kept for no longer than is necessary.
- handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction, or damage.
- used responsibly ensuring compliance with the principles of the UK GDPR.

c. Legal basis

The UK GDPR and the DPA set out the ways in which personal data can be lawfully processed (Article 6). All processing of personal data must be on the basis of at least one of the following:

- **consent:** the individual has given clear consent for you to process their personal data for a specific purpose
- **contract:** the processing is necessary for a contract you have with the individual, or because they have asked you to take specific steps before entering into a contract
- **legal obligation**: the processing is necessary for you to comply with the law (not including contractual obligations),
- **vital interests**: the processing is necessary to protect someone's life
- **public task:** the processing is necessary for you to perform a task in the public interest or for your official functions, and the task or function has a clear basis in law
- **legitimate interests**: the processing is necessary for your legitimate interests or the legitimate interests of a third party, unless there is a good reason to protect the individual's personal data which overrides those legitimate interests. (This cannot apply if you are a public authority processing data to perform your official tasks).

Under the UK GDPR, health data is defined as special category data (that is data that requires additional protections due to its sensitivity). For this type of data to be processed a further condition must be met in addition to one of the legal bases set out above (Article 9). These conditions could be:

- explicit consent
- employment, social security, and social protection (if authorised by law)
- vital interests
- not-for-profit bodies
- made public by the data subject
- legal claims or judicial acts

- reasons of substantial public interest (with a basis in law)
- health or social care (with a basis in law)
- public health (with a basis in law)
- archiving, research, and statistics (with a basis in law)

d. Common Law

Alongside data protection legislation, the common law duty of confidentiality also applies to the use of confidential patient information.

The general position is that if information is given in circumstances where a duty of confidence applies, that information cannot normally be disclosed without the information provider's consent.

The three circumstances where disclosure of confidential patient information is permitted are:

- where the individual to whom the information relates has consented
- where disclosure is necessary to safeguard the individual, or others, or is in the public interest
- where there is a statutory basis for disclosing the information or a legal duty (such as a court order) to do so

e. Monitoring

In the UK, the Information Commissioner upholds information rights as the independent regulator dealing with the Data Protection Act 2018 and the UK General Data Protection Regulation. The Information Commissioner investigates complaints of breaches of data law and can issue fines where complaints are upheld.

The Information Commissioner also provides guidance on the operation of the DPA and UK GDPR which can be found on its website.

In England, the National Data Guardian for Health and Social Care is a statutory role established to advise and challenge the health and care system to help ensure that an individual's confidential patient information is safeguarded securely and used properly.

The previous National Data Guardian, Dame Fiona Caldicott, also established the Caldicott Principles which inform the use of confidential patient information in the health and care system. These principles are:

- justify the purpose(s) for using confidential information
- don't use patient identifiable information unless it is necessary
- use the minimum necessary patient-identifiable information
- access to patient identifiable information should be on a strict need-to-know basis

- everyone with access to patient identifiable information should be aware of their responsibilities
- understand and comply with the law
- the duty to share information can be as important as the duty to protect patient confidentiality

All NHS organisations and local authorities that provide social services in England must have a Caldicott Guardian to uphold these principles and protect the confidentiality of people's health and care information, making sure it is used properly.

All four nations have chosen to have Caldicott Guardians. These are represented by the UK Caldicott Guardian Council, which is a sub-group of the National Data Guardian's Panel.

Annex 3 – Survey questions

Welcome: Instructions

1. Please complete this survey to provide examples on barriers you have experienced regarding health data sharing.
2. Please submit the survey after each example you have provided. You can provide multiple examples by submitting the survey multiple times.
3. In case you have experienced a barrier regarding health data sharing which is not listed here you can add one using the 'other' option.
4. There are no mandatory questions. If you do not wish to provide a response, please simply move on to the next question.
5. For each example, we are interested to know specific details (e.g., costs and implications) and your recommendations to mitigate and/or resolve these barriers. Please be as specific as possible.

Your responses will be used to inform the development of the European Health Data Space and the corresponding legislation.

If you have any questions, please contact: TEHDAS.sciensano@sciensano.be

Page 1: Background info

- Name: _____
- Surname:_____
- Email address:_____
- Organisation/institute:_____
- Country:_____
- Job Title: _____
- Job Type (***Choose all that apply***: Researcher, Policy maker, Data holder, Data consumer, Data controller, Other: ___)
- Select a barrier you would like to provide an example for: (***Choose one of the following answers***)
    a. There are differences in governance and health data systems in Europe. Please map the governance and health data systems in your country.
    b. There is no common European interpretation of what constitutes 'sufficient anonymisation' to transform personal data to non-personal data.
    c. There is no common European interpretation of what constitutes 'pseudonymisation'
    d. There is no common European interpretation of what is, and what is not, 'secondary use' of data.
    e. European countries have national laws/ rules around health and research data in addition to GDPR.
    f. European countries have the ability to set their own derogations under the GDPR. This lack of harmonisation creates additional barriers.
    g. European countries have different preferences as to the choice of legal basis for processing under the GDRP. This created barriers to cross-border collaboration and data sharing.

h. Health data is considered sensitive data e.g., special category data under GDPR and is treated differently from other types of data when it comes to health data ethics, management and use.
i. No standardised data sharing agreements exist for products developed by private sector providers using public sector health data to (a) facilitate safe data sharing and (b) protect taxpayers' investment
j. Across Europe, different taxonomy and ontology codes are used to label the same health condition, making comparisons between data sets challenging
k. Poor data management procedures reduce the ability to reuse data.

If you selected barrier a. There are differences in governance and health data systems in Europe.

Section 1: The structure of the health data system in your country
Please outline the structure of the health data system and name the health data organisations that form this system in your country.
For each health data organisation please indicate:
(a) Whether they are: i. Data processor; ii. Data controller; iii. Both, iv. Neither
(b) Where the organisations' data comes from?
(c) What type of data does the organisation have? E.g., genomic, health records, audit etc.

Section 2: The governance of the overarching health data system in your country
Please outline the governance of the overarching health data system in your country
1) What is the process to gain access to the data in your health data system?

2) What is the process for monitoring and providing access to data in your health data system?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier b. There is no common European interpretation of what constitutes 'sufficient anonymisation'

Does your country have national level guidance on anonymisation of health data and/ or data? Do you use it?

In your opinion, how do you define 'sufficient anonymisation'?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____

- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier c. There is no common European interpretation of what constitutes 'pseudonymisation'

Does your country have national level guidance on pseudonymisation of health data and/ or data? Do you use it?

In your opinion, how do you define 'pseudonymisation'?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier d. There is no common European interpretation of what is, and what is not, 'secondary use' of data

Does your country have national level guidance on what is, and what is not 'secondary use' of health data and/ or data. Do you use it?

In your opinion, how do you define 'secondary use' of health data and/ or data?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier e. European countries have national laws/ rules around health and research data in addition to GDPR

What national legislation or rules in regard to health and research data does your country have in addition to the GDPR?

Is there conflict between the GDPR and other national legislation or rules? If yes, how do you manage this conflict?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier f. European countries have the ability to set their own derogations under the GDPR

What GDPR derogations exist in your country for scientific research purposes and health (Article 89)?

Does the lack of harmonization of derogations impact your work? If yes, please give an example.

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier g. European countries have different preferences as to the choice of legal basis for processing under the GDRP

Does your country have guidance on the choice of legal basis for (a) scientific research purposes (b) health? Please could you share a link to this guidance.

Do the different preferences of Member States, in relation to the legal basis for data sharing, impact your work?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier h. Attitudes to use of health data (GDPR special category data) can cause risk-averse behaviours

Have you experienced difficulties accessing or sharing health data despite having a clear legal basis for doing so? How did you manage this situation?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier i. No standardised data sharing agreements exist for private sector use of public sector data

Does your country/ organisation have any examples of standardised agreement templates that would be appropriate for public and private partnerships? Please share a link(s) if possible

In your opinion, what conditions should be included in public private partnership agreements to ensure that the public data is secure?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier j. Across Europe, different taxonomy and ontology codes are used to label the same health condition

Do you have examples of initiatives or methods to support interoperability between multiple data sets? Please describe.

What recommendations would you suggest to improve interoperability?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?

If you selected barrier k. Poor data management practices reduce the ability to reuse data

Do you have examples of good practice in data management to support the identification and access to data sets?

Do you have examples of good practice in data management to support the reuse of data?

Please provide a specific example of a data sharing process you have experienced regarding the selected barrier, by answering the following questions
Describe the impact of the barrier on data sharing

What is the level of data sharing?
(***Choose one of the following***: Regional, National, European, International, Other:__)

What is the country(ies) and organisation(s) of data origin? _____
What is the country(ies) and organisation(s) in data receipt? _____

What type of data was used?
(**Check all that apply**: Personal data, Non-personal data, Anonymised data, Pseudonymised data, Aggregated data, Individual level data, Other:_____)

What type of data set was used? (indicate whether it was public data, hospital data (EHR), genomic data, etc)

What is the role of each actor in this process? (as applicable)

- Data subject:_____
- Researcher:_____
- Policy maker:_____
- Data processor:_____
- Data controller:_____
- Other:_____

Have you overcome this barrier? If yes, how? (Please elaborate and provide references if applicable)

What are the costs (financial, resource, other) caused by this barrier and its mitigations?

What are the ethical implications of this data sharing process?

Please outline any other best practice solutions or recommendations to mitigate/ resolve this data sharing barrier?