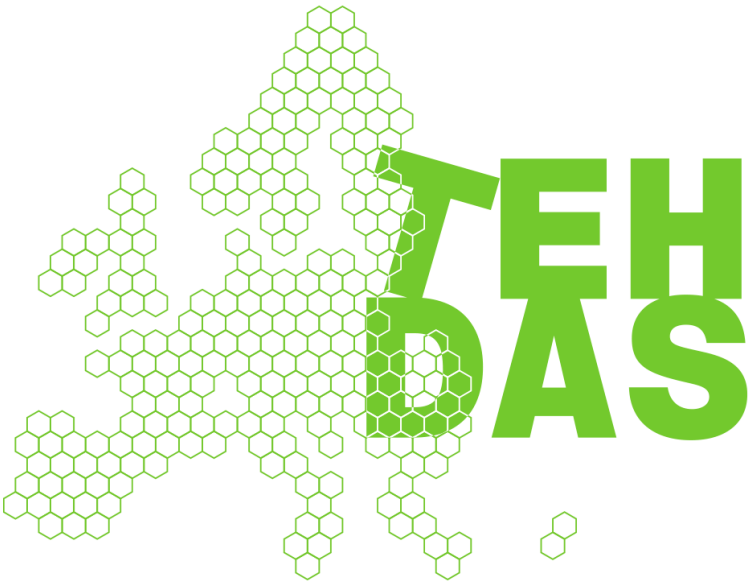

Towards European Health Data Space

Summary of Milestone 5.1 & 5.2

Summary of results: case studies on barriers to cross-border sharing of health data for secondary use

Authors: Linda Abboud, Shona Cosgrove, Irini Kesisoglou, Rosie Richards, Flavio Soares, Cátia Pinto, Petronille Bogaert, Sarion Bowers

28 September 2021

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

Accepted in Project Steering Group 28 September 2021.

Disclaimer

The content of this deliverable represents the views of the author(s) only and is his/her/their sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use of its contents.

Contents

Annex A as a separate document.

1 Introduction

TEHDAS, the Joint Action Towards the European Health Data Space, helps EU Member States, associated countries and the European Commission to develop and promote concepts for the secondary use of health data to benefit public health and health research and innovation in Europe. The results of TEHDAS will support the European Commission in building a European Health Data Space by developing the principles for data sharing, providing different proposals and governance models for secondary use of health data.

TEHDAS consists of eight work packages including a specific work package on 'Sharing Health Data'. The overall aim of the Sharing Health Data work package is to provide options for the operational framework and governance models for the exchange and secondary use of health data between European countries, respecting the principles of transparency, trust, FAIRness, citizen empowerment and common good.

As part of the Sharing Health Data work package, TEHDAS completed research and analysis to define and develop the evidence base for secondary use of health data by identifying data sharing practices across Europe. The work can be divided into three work strands:

1. **Literature Review**: A focused literature review 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 development:** The development of a framework to provide an overview of the current barriers to cross-border sharing of health data in EU Member States and associated countries, based on the literature review and prioritisation analysis.
3. **Case Studies:** The development of case studies in collaboration with stakeholders from EU Member States and associated countries to create the evidence base for the secondary use of health data from the perspective of users, researchers and policymakers.

2 Literature review

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 research question which guided the review was: 'What are the barriers to cross-border sharing of health data for secondary use, for non-personal health data and personal health data under the GDPR?'.

The rapid review identified a list of 90 barriers and 73 enablers for data sharing across Europe which were compiled into a framework and organised into specific themes. The specific themes of the barriers and enablers identified in the literature review can be categorised under the following titles:

- **Legal:** including semantics, legal frameworks, and national interpretations of GDPR
- **Data:** including data management, data quality, data interoperability, data monitoring and analysis
- **Trust and transparency:** including political, social and organisational factors and citizens' engagement
- **Infrastructure:** including the governance structure of the health data system and access to data
- **Resource:** including human, financial and technical resources
- **Ethical aspects:** including patient's informed consent, privacy protection to patients and de-anonymisation aspects

3 Developing a framework: analysis of priority barriers

Based on the literature review findings a framework was developed and a list of 20 barriers were identified for further analysis. The list included barriers representing the themes: Legal, Data, Trust and Transparency, and Infrastructure. The remaining themes (resource and ethical aspects) were incorporated in the case study template as overarching themes. This list was further refined through a prioritisation exercise involving 18 countries' representatives to TEHDAS. The final list of barriers is presented in table 1.

[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

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 Legal
B	There is no common European interpretation of what constitutes 'sufficient anonymisation' to transform personal data to non-personal data.	Legal
C	There is no common European interpretation of what constitutes 'pseudonymisation'.	Legal
D	There is no common European interpretation of what is, and what is not, 'secondary use' of data.	Legal
E	European countries have national legislation/rules around health and research data in addition to the GDPR.	Legal
F	European countries have the ability to set their own derogations under the GDPR. This lack of harmonisation can create additional barriers.	Legal
G	European countries have different preferences as to the choice of legal basis for processing under the GDPR. This creates barriers to cross-border collaboration and data sharing.	Legal
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.	Data
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.	Trust and Transparency
J	Across Europe, different taxonomy and ontology codes are used to label the same health condition, making comparisons between data sets challenging.	Data
K	Poor data management procedures reduce the ability to reuse data.	Data

4 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 specific examples (case studies) from experts, institutes and/or projects within EU Member States or associated countries. The aim of the case studies is to substantiate barriers and identify potential best practices and solutions to improve sharing data between European countries. The case studies also aim to explore the purpose, requirements, and type of data, focusing on scientific research and innovation and policy making for public health purposes.

5 Summary

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

Figure 1: Number of entries per barrier

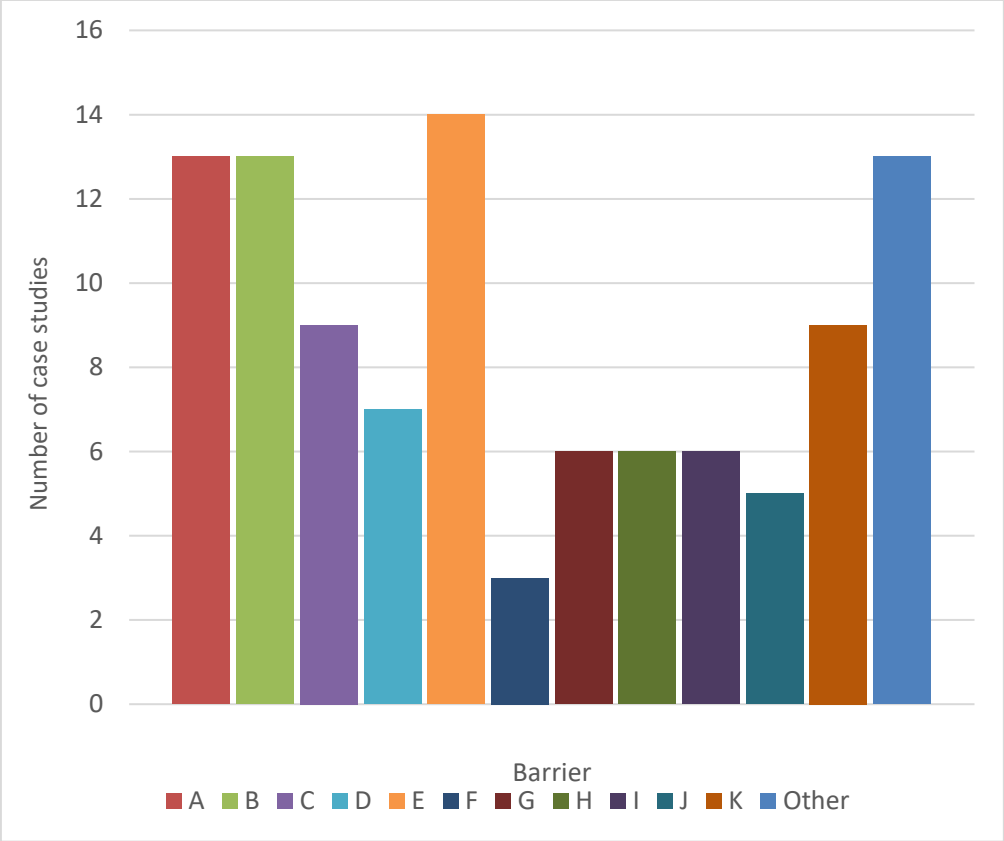

Figure 1: Number of entries per barrier shows the distribution of case study submissions per barrier. The barriers F and J had the lowest responses (≤ 5 responses). "Other" represents case studies for barriers that participants perceived as a priority but were not included in the predefined list of 11 barriers.

Figure 2: Countries that contributed case studies

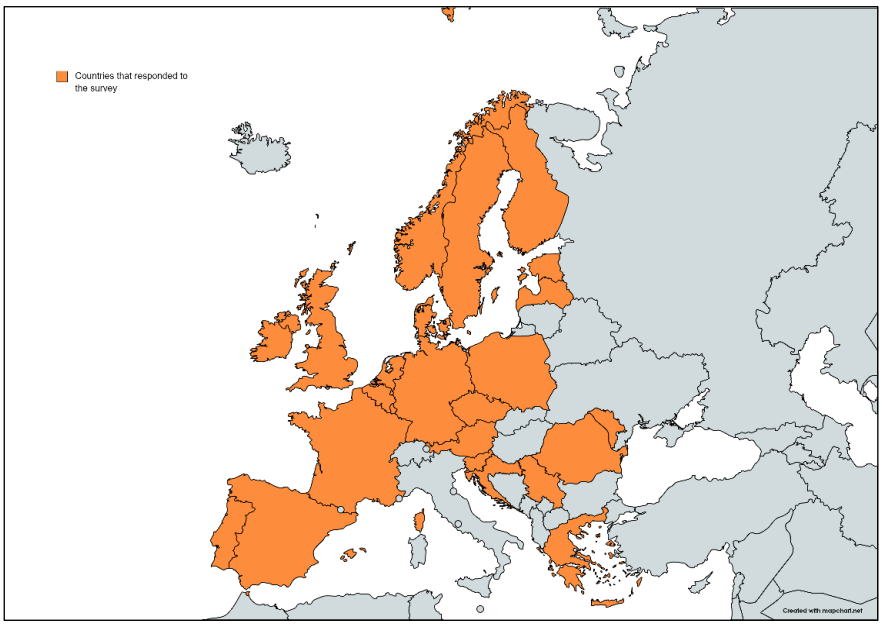

Figure 2: Countries that contributed cases studies provides the overview of the distribution of countries that responded to the survey. A total of 23 countries provided at least one case study, representing a wide coverage of European countries.

Figure 3: Case studies grouped according to sector

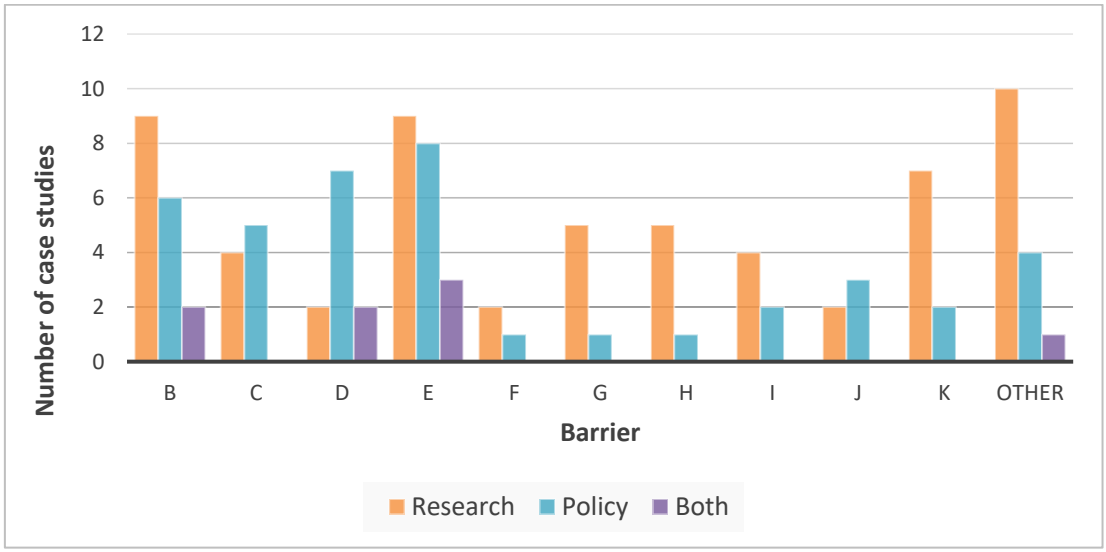

Figure 3 - Case studies grouped according to sector shows the number of case studies submitted by users from research and/or policy perspective. The 'Both' category represents case studies where the contributor selected both 'researcher' and 'policy maker' in the job type field. Barrier A, governance and health data management systems in Europe, is not included in Figure 3 - Case studies grouped according to sector as it provided a description of the health data management system in a country, rather than an experience with a barrier.

5.1 Barrier A: There are differences in governance and health data management systems in Europe.

A cross-section of participating European countries was asked to map the governance and health data management systems in their countries. Ten countries (Austria, Belgium, Denmark, Estonia, Finland, Greece, Ireland, Moldova, Sweden, and the UK) provided a national-level snapshot. This range of countries provided a good representation of countries across the European Union and in wider Europe.

The countries described various health data management models ranging from centralised, decentralised, and federated systems. Significant differences in governance and national specificities were evidenced across the submissions. Countries have taken very different approaches in the development of their health data management systems. Stakeholders 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. This summary report is complemented by Annex A which presents health data management profiles for participating countries.

5.2 Barrier B: There is no common European interpretation of what constitutes 'sufficient anonymisation' to transform personal data to non-personal data.

Across the board, researchers and policymakers reported a lack of guidance on anonymisation at national and international level as a key barrier to data sharing. In particular stakeholders identified a lack of clarity between "absolute" and "relative" anonymisation as a key issue. Stakeholders also identified a lack of anonymisation guidance for medical images, genomic data, longitudinal data, and rare diseases. Finally, other commonly cited issues were how to define the parameters for re-identification.

Impacts	• Interpretation of applicable methods for anonymisation varies significantly among regional, national and European authorities, causing internal interoperability issues. • Risk-averse behaviours due to lack of clarity. Some stakeholders reported that they treat all data as personal data due to this lack of clarity. • Stakeholders stated that some countries apply a stricter definition of 'sufficient anonymisation' which further limits the sharing of data for research on the basis that the individual could potentially be traced and re-identified e.g. due to the rarity of their illness. • Speed of innovation is reduced or impeded. • Unclear public communication around health data use.

	• Uncertainties about how and if certain types of personal health data can be accessed. Secondary impacts include delays and financial costs. • 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. • 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.
Recommendations	• Applying anonymisation methods including grouping, noise adding and randomisation. • Synthetisation of the data (following anonymisation). • Applying defined automated anonymisation methods and tests. • Utilising AI based anonymisation software tools. • Employing approved anonymized public use files such as the Canadian ICES. • Creating mechanisms for international data sharing based on safeguards other than Standard Contractual Clauses in isolation. • Finance controlled access repositories. • Identify suitable access models, such as broad consent from patients to allow for future use of datasets. • Align with the US Food and Drug Administration standards. • Recognise the US's PHUSE anonymisation methods and principles. • Develop and implement anonymisation guidance at a global level. • Create clear guidance on when data can be accepted as anonymised vs pseudo-anonymised for health data and imaging.

Stakeholders' responses showed that determining when data can be considered sufficiently anonymised is viewed slightly differently by each European country and at national and

regional levels. However, stakeholders were united in calling for compatible anonymisation guidance at all levels as well as for checklists to provide certainty on the processes to anonymise data.

5.3 Barrier C: There is no common European interpretation of what constitutes 'pseudonymisation'.

Responses consistently identified a lack of European or national level guidance on the pseudonymisation of health data or data as the primary issue. Other issues identified were caused by this lack of guidance. Three main sub-issues were identified. Firstly, the different approaches taken by data controllers, data protection officials and ethical review boards towards pseudonymisation at regional, national and European level, creating interoperability issues and hindering data sharing. Secondly, a lack of consensus on the necessary degrees of separation between the key to the data user in order for the data to be classed as unidentifiable. Finally, a lack of clarity around when it is in the public interest to provide pseudonymised data as opposed to anonymised data for research.

Impacts	• Data pools may not be interoperable because of different standards and methods of pseudonymisation. • Data are sometimes required to be sent to a third party in order to align the format and combine two separate data sets. This can require 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 rare diseases sector. • The application of safeguards adds significant cost and resource requirements. • The need to create individual solutions for each project requires time and resources. • Financial costs are increased due to repeated processes and the time and resource commitments required to establish data sharing agreements and to prepare the data. • In some countries only aggregated data, and not pseudonymised data can be shared for secondary use and research purposes.
Recommendations	• Guidance to harmonise pseudonymisation methodologies. • Align the assessment tools used by data protection officials.

	• A Code of Conduct to achieve sufficiently robustly pseudonymised personal data in the public interest without data subjects' explicit consent. • Clear guidance around third parties holding pseudonymisation keys and distance from that key. • Pseudonymised data could be processed and analysed in a special "safety room", from which only aggregated data and outputs can be shared. • A law or code of conduct that provides the opportunity to share data for studying rare diseases. • Establishing a pseudonymisation expert group at an EU level.

There was a resounding call from stakeholders for guidance and clarity on the methodology for and interpretation of pseudonymisation. However, at the same time there was a clear steer from respondents that any guidance would need to avoid a one-size-fits-all approach.

5.4 Barrier D: There is no common European interpretation of what is, and what is not, 'secondary use' of data.

Stakeholders pointed to the lack of national and European level definition of the secondary use of health data as a barrier to research. A particular issue stakeholders asked to be addressed as a priority is a lack of clarity around consent, especially broad consent.

Impacts	• Reluctance from organisations to access data due to the lack of clarity on the accessibility procedure. • Data coupling to include socio-economic and behavioural population data is often extremely difficult, costly or time consuming. • Some stakeholders reported that they treat all data as secondary data as a preventative measure. • Difficulties for sponsors and ethical review boards to determine if consent has been given for subsequent uses of clinical study data. • If sponsors cannot conduct analysis of prior trial data without anonymisation, it may render the research useless.

	• If sponsors have to reconsent 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. • Certain research studies cannot be conducted due to unclear definition of what is secondary use of health data, its purpose and whether it is compatible with what is allowed.
Recommendations	• Legislation specific to health information that clearly outlines the rules for the collection, use and sharing of personal health information for secondary use. • A specific consent model to enable the processing of personal health information. • Data use which adheres to the FAIR principles (Findable, Accessible, Interoperable, Reusable). • Clear EU-wide definition and guidance on what is secondary use of health data.

Stakeholders were united in calling for national and European action to provide clarity and give guidance to controllers and processors on the secondary use of health data especially for scientific research purposes. Stakeholders called for the EU to support Members States to put in place structures allowing for secondary use of health data for research and to support the industry in processing health data to realise better patient outcomes and healthcare systems.

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

Stakeholders highlighted the difference in interpretation of the GDPR rules in Europe and the existence of additional national rules. Stakeholders report that these additional national rules can cause complications in the exchange of health data for secondary use between Member States and associated countries.

Impacts	• A lack of clear rules for the use of health data for research, particularly if the researcher does not work for the data holder, causes a broad range of issues regarding availability and conditions for the secondary use of health data. • 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 holders.

	• Differing approaches to the use of consent exacerbate difficulties to obtain consent retrospectively for secondary use of health data. • A disproportionate "stacking" of multiple overlapping safeguards required by different jurisdictions. • Where cross-border data processing takes place, the different national legislations have to be applied at the same time. This leads to different processing requirements, which hampers joint research projects. • There is a conflict between outdated national protection laws and GDPR limits sharing of data. The existence of overlapping acts at EU and national level has led to differences in interpretation and applications. • Health data is underutilised as a resource for secondary use • There is an impact on the reuse of data and the motivation of industry players to invest in databases to allow the efficient sharing of data for secondary use. • Risk-averse behaviour leading to reliance on consent, including when GDPR does not require it, is an obstacle for scientific research.
Recommendations	• Tangible and practical action plan for international data sharing (not just legal). • Greater harmonisation and consistency across health ministries, ethics committees and hospitals. • Sharing parameters instead of sharing data will be a game changer in AI applications in the health sector. • Harmonisation should not only apply to data from healthcare but to all health-related data (including molecular data). • To harmonise, update or amend national legislation to remove any conflict across different acts. • Consensus on how to interpret GDPR in relation to national laws. • To move research activities to countries outside of the EEA. • Health specific legislation clearly outlining the rules for the collection, use and sharing of personal health information. • A specific consent model precisely defining the circumstances whereby explicit or implied consent is required and where there

	is a legal basis for the use of personal health information without consent. • Harmonised EU rules on health data encompassing intellectual property stakes and the ethical valuation of data. • A European framework enabling industry, hospitals and researchers to process patient data. • Logging and tracking digital access to patient clinical files, automatically reporting activity to the patient.

5.6 Barrier F: European countries have the ability to set their own derogations under the GDPR.

Stakeholders reported two main issues. Firstly, that guidance from national data protection authorities and the European Data Protection Board on how GDPR should be understood, interpreted and applied in various circumstances is still forthcoming. Secondly, where different rules apply in different countries, cross-border consortia are hampered.

Impacts	• Difficulties in creating an approach to the processing of personal data which would be legally compliant across Europe. Cross-border consortia are hampered. This challenge is made more acute with regards to the fact that the medical data at the heart of this project is 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 becomes hampered as data subjects' rights can impair or make the research project impossible where no derogation was foreseen under national law. • Delays due to awaiting clarification on the GDPR positions as adopted by individual Member States and associated countries for guidance from regulatory and professional bodies. • Organisations risk severe delays to research projects. • Research projects may be hampered as some researchers prefer not to use personal data in order not to fall under GDPR due to the delays that this could cause.

	• Broad consent in some countries does not include the export of pseudonymised information, especially from large older cohorts, limiting access and sharing of data cross borders. • There is uncertainty over the lawfulness of using consent when it should also cover secondary use, especially when conducted cross-border by a separate research organisation. Consent accepted in one country with a certain broadness, may validate subsequent recipients and/or collection context in one European country but not in another country. This could create massive disruptions in data sharing.
Recommendations	• Ensure that the governance and infrastructure of the data platform are both centralized and federated (allowing data to stay behind institutional firewalls). • The European Health Data Space (EHDS) could make use of Art. 89 but also Art. 23 of the GDPR to create harmonisation within the EHDS and provide derogations important for health research. • Creating an EU law for the EHDS, where Union law could be the source for creating the basis for certain types of processing or derogations (such as research in the public interest; processing of health and genetic data without consent; derogation from data subjects' rights; research use of data and the corresponding safeguards).

It is important that national competence and the ability of Member States and associated countries to set their own derogations is respected while at the same time Europe works together to ensure data sharing practices are as streamlined and harmonised as possible. Stakeholders have principally called for clarity on national derogations. It is still the case that many European countries' data protection legislation is available only in their national language adding further confusion over what derogations do or do not apply in Member States and associated countries and how to share data across consortium with differing rules. This is particularly true for health data sharing which is a 'special category data' and even more so for areas like genomics where significant and varying derogations apply by country.

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

Stakeholders reported two main issues. Firstly, a lack of guidance and agreement on the choice of legal basis for data processing. For example, due to the different preferences of legal basis, data are collected or made available using consent in one country and using public interest or legitimate interest in another country. Secondly, there are no regulations on how to archive video data, and other personalised content linked to the video data are

missing. It is unclear if and how secondary data analysis of narrative, ethnographic, video data is allowed.

Impacts	• The lack of guidance on the choice of legal basis results in research institutions interpreting GDPR in relation to each separate contract. • The lack of legal basis for certain types of data, such as audio-visual data, creates confusion on the allowed use and analysis of such data. • A controller may have to apply different legal bases for the same processing where data are collected from different countries and/ or sources. • Different legal bases may apply to different datasets as well as to individual data types. • The data sharing situation is difficult to explain to data subjects. • A controller who is obliged to process under consent, but where the 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. Some controllers e.g., private entities may not have any possibility at all to process such data. The same is true for some data types such as genomics where consent is required for processing and no derogation is foreseen in the law. • The rights under the GDPR depend on the choice of the legal basis. Where certain rights apply under one controller, they may not apply under another, or they may only apply to part of the data. • Different preferences of conditions for data sharing among Member States and associated countries hampers or could hamper successful implementation of transnational research projects and pan-European initiatives (joint calls for COVID-19 research, 1+ Million genome initiative etc.). • Serious delays have been experienced in research as each institution has a different approach to contracting for projects. • Human resources and financial costs are required to maintain and monitor the individual contracts and to fulfil the requirements of the GDPR.
Recommendations	• A controller who is obliged to process under consent, but where the 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. Some controllers e.g., private entities may not have any possibility at all to process such data. The same is

	true for some data types such as genomics where consent is required for processing and no derogation is foreseen in the law. • The rights under the GDPR depend on the choice of the legal basis. Where certain rights apply under one controller, they may not apply under another, or they may only apply to part of the data. • Different preferences of conditions for data sharing among Member States and associated countries hampers or could hamper successful implementation of transnational research projects and pan-European initiatives (joint calls for COVID-19 research, 1+ Million genome initiative etc.). • Serious delays have been experienced in research as each institution has a different approach to contracting for projects. • Human resources and financial costs are required to maintain and monitor the individual contracts and to fulfil the requirements of the GDPR. • A united approach on using public benefit as legal basis for health data sharing across all European countries. • Harmonisation of European and national level rules on how to archive video data and other personalised content linked to the video data.

Stakeholders called for clarity and a united approach to the use of a legal basis at European, national and regional level. They welcomed the European Data Protection Board Opinion of 23 January 2019, which stated that consent is not the only legal basis for research purposes and other legal bases should be supported. However, over 2 years have passed, and according to stakeholders this change has not yet been implemented at national level. Stakeholders state that they continue to face the same issue with consent often being seen as the only available legal basis.

5.8 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.

Stakeholders reported overly risk-averse behaviours in applying the GDPR, resulting in requests for health data being rejected including in instances when legal and ethical approvals were in place. Some Member States and associated countries reported that hospitals in their countries cannot share health data with private companies, including when such companies may process the data in the public interest. Stakeholders also highlighted the need for clarity on the use of consent for secondary use and particular challenges around the sharing of genomic data.

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. • The need to move certain operations outside EU countries and the associated financial and resource related costs. • A lack of clarity on the legal basis for genomic data use.
Recommendations	• Clarity on how to treat special category (health) data when it comes to health data ethics, management and use. • Development of a federated infrastructure with a distributed architecture removing the need for any individual data sharing. • Standardised guidelines to facilitate agreement on which data or metadata can be published and which cannot. • Awareness-raising and best practices to ensure data holders do not display overly risk-averse behaviours which go beyond the prevailing ethical and legal frameworks. • Establishing agreements on FAIR-based, open access metadata. (Findable, Accessible, Interoperable, Reusable). • Clear information regarding the use of human body material, including health data in the framework of genomic research. • Organisations dealing with vast amounts of sensible health data should invest in hiring or subcontracting Data Protection Officers.

A number of the stakeholders stated that they have not been able to mitigate or overcome this barrier and had been forced to move processing outside of the EU or use consent as the prevailing legal basis going against the advice of the European Data Protection Board. Clarity and awareness raising were seen as key to addressing the principal issue, a culture of overly risk-averse behaviour in relation to sharing special category (health) data.

5.9 Barrier 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.

Under this barrier, stakeholders highlight the importance of establishing rules and guidelines for equal access to health data for the public and private sectors. This is important to ensure research and innovation development.

Impacts	• If the provisions in the data sharing agreement are disproportionate 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 stakeholders have equal, regulated access to health data so that fair competition for best solutions can take place. • The medical technology industry has no access to the research data center. Private organisations are unable to access billing and health data from EHR from insured individuals. • Due to lack of standards, to make private public collaborations possible, a tailor-made approach must be implemented to any given situation. • Loss of the development opportunities • Costly and long process of obtaining data for secondary use. • Limitation of scope of shared data because of lack of confidence at disclosing party. • There is no standard format for applying for data resulting in duplication of work and confusion. • High financial impact as the private organisations that own data ask for significant amounts of money for data that is being collected using public money. • Cloud computing services owned by a U.S. 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.

Recommendations	• A uniform jurisprudence. Standards must not be set unilaterally for the national healthcare market, but on the basis of international standards together with the industry. • Implement a well-developed and widely accepted framework (standard policies) for the secondary use of data. • Guidelines on proportionate measures may be helpful. • Establish an agreement on common rules for collaboration on quality registers between association of local authorities and the pharmaceutical industry. • FAIR criteria (Findable, Accessible, Interoperable, Reusable) should be incorporated into agreements, especially the accessibility criteria. • A rule that all data that are being collected using public money, routinely or not, should be available for free from month 18 after collection onwards.

5.10 Barrier J: Across Europe, different taxonomy and ontology codes are used to label the same health condition, making comparisons between data sets challenging.

Stakeholders reported a number of issues regarding semantic interoperability. The use of different terminologies creates significant problems for combining datasets, analysis and common understanding. In the rare disease sector this has caused delayed responses to urgent requests and has often led to misunderstandings and diagnostic errors.

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 (MPS) and hospital information systems (HIS).
Recommendations	• Synchronisation of transformation to the WHO International Classification of Diseases 11th edition. • Roll out adoption of SNOMED CT-Orphacode mapping tools to ensure smooth incorporation of Orphacodes into all EU national EHR systems. • Develop national protocols for handling of rare diseases not in the SNOMED CT-Orphacode mapping set and undiagnosed/undiagnosable rare diseases, with modifications to the national reference set as necessary. • Incorporate new terms in SNOMED for undiagnosed and undiagnosable rare diseases. • Comprehensive and consistent use of a classification or taxonomy/ontology like ICD-10 or SNOMED CT for all documents and messages produced in the MPS/HIS.Comprehensive mapping between ICD-10 and SNOMED CT. • A national infrastructure and standardisation. A national data entry point that can support and collect data of different levels. • Federated analysis where there is no need to share patient level data. • Adopt the EMA's one coding system for medicines. The project is currently in development. • To promote the use of ontology servers, open published tools and tool sharing.

Stakeholders recommend promoting the harmonised use of ontology servers and open published tools and tool sharing for example HPO, SNOMED CT, ICD-11. Stakeholders believe that this will improve semantic interoperability in all different health data types, from EHR sources to rare diseases taxonomy.

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

Stakeholders report that poor health data management is a barrier to health data exchange for secondary use between institutes within the same country and across Europe. The lack of adherence to the FAIR principles (Findable, Accessible, Interoperable, Reusable) is reported to be one of the sources of the issue.

Impacts	• Poor data management causes loss of value of information being generated because of missing or inconsistent data entry. • Slow access to data, which then causes delays and difficulty ascertaining where to access the data. • Poor management makes the reuse of data time demanding. • A lack of transparent health research project approval process cause confusion. • Slow processes for data integration for large projects. There is no developed system for federated data integration for public data, hospital data (EHR) and, or 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 pre-GDPR data 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.
Recommendations	• Policies which provide data controllers with consistent practices or virtual secure analysis environments. • Develop an identification system at EU level, similar to the European Commission research funded projects Participant Identification Code, which could help to streamline the data access request across Europe for trusted organisations.

	• Use aggregated data between regions. In this way, the previously sensitive data becomes less sensitive and can easily be shared between peers for reuse. • Suppliers could adopt the clinical information model, fully enabling the reuse of data from the electronic health records. • Make metadata descriptions to data sets mandatory, with a standardised set of what should be described at the bare minimum. • National list of health data sources and data controllers, with a research focus and with good metadata descriptions for each data source. • Consistent policies which define access protocols. • Designated individuals to facilitate research access. • Secure environments which allow researcher access. • National Research Ethics Committees. • Single point of access to several datasets. • Federated data integration for public data. • Standardised terminologies and/or coding systems used at point-of-care if possible. • Standardised data dictionary with definitions which data controllers must abide by. • OpenData initiatives and portals. • Additional guidance on data management and training on epistemology and technology of research, and investment in robust IT infrastructure. • Better automated anonymisation on many dimensions to prevent data leakage or data reconstruction. • EHR-systems could provide a user-friendly interface which supports and facilitates the registration of (structured) data according to the workflow. Only then can the transition to a 'circular-healthcare-data' be achieved.

5.12 Other barriers

Barrier and impact	• There is no common definition or understanding of roles in data processing i.e., who is controller, joint controller, processor or data recipient. Member States and associated countries have different interpretations. • There is still a lack of clarity regarding to what extent data that was collected prior to GDPR coming into force can still be used.
Recommendations	• Clear definitions and understanding what kind of data are being processed, for what purposes and by whom. • Establish a joint European data sharing formula (joint controller/ data processor agreement) that all countries could agree upon, would interpret in the same way and use the same formulas when initiating collaborations. • Legal certainty on the processing of health data, and careful consideration of the possible retroactive effects it may have that would cause unreasonable impacts.

6 Conclusion

This report summarises the case studies and perspectives provided by researchers and policy makers on the barriers they experience to cross-border data sharing of non-personal and personal data. More specifically, the report provides an overview of the impact of the barriers and proposed recommendations provided by the respondents.

The case studies show that researchers and policy makers experience a variety of barriers in reusing and cross-border sharing of health data. While the summary section does not distinguish between the experiences of researchers and policy makers, figure 4 shows a higher response rate from researchers to certain barriers, compared to policy makers, and vice-versa.

For example, policy makers are concerned with the lack of common interpretation of secondary use of data as it inhibits the ability to use health information beyond the direct care of patients. Whereas researchers are concerned with the lack of common interpretation because it inhibits further processing of health data which has already been “collected”.

However, the number of responses per barrier is not large enough to make a distinction between barriers that are more important for research or for policy makers. Therefore, it makes sense to continue to treat these perspectives together, whilst acknowledging the nuances in their perspectives.

The next phase of this work is to draw on the case studies and evidence provided by stakeholders to inform the development of options to mitigations and resolve to the barriers identified. The final report and options will inform wider TEHDAS work streams including the development of guidelines for Member States when developing national legislation on secondary use of health data and European Health Data Space architecture and governance options. To achieve this, we will:

- Further analyse the case studies and summarise on a multinational level the conclusions for the respective user categories in scope (researchers and policymakers).
- Develop policy options to mitigate and resolve the identified barriers in the EHDS.
- Develop policy options to support best practice in data sharing for secondary use across Europe.
- Conduct additional stakeholder consultations to test the proposed policy options.
- Continue to input to and support the wider Sharing Health Data work package and the Joint Action to develop policy on the European Health Data Space, ensuring consistency and knowledge transfer between projects.

If you would like to learn more about the project or participate in the consultation process, please register your interest by contacting: TEHDAS.sciensano@sciensano.be.