



# No time to lose: Building a data strategy for the European Reference Networks

A EURORDIS contribution

April 2020



# Contents

Executive summary .....	2
1. Background and objective .....	5
2. Why we need to act now .....	7
3. What are the health data needs of the ERNs? .....	13
4. Defining and delivering a health data strategy for the ERNs.....	15
1. Start the conversation to develop a shared vision with all stakeholders .....	17
2 Set up an implementation framework.....	22
5. Conclusion.....	25
Annex I. Definitions and De-identification Terms .....	26
Bibliography .....	28
Additional References.....	29

# Executive summary

The [European Reference Networks](#) (ERNs) are expected to transform the wealth of health data and collective knowledge on rare diseases scattered throughout Europe into improvements in the delivery of highly specialised care and a striving rare disease research landscape.

These Networks are a groundbreaking new structure that represent a unique opportunity for the innovative use of health data across borders to improve the lives of people living with a rare disease. At the same time, the large majority of people living with a rare disease are willing to share their health data to advance care and research, as long as this is done in a secure manner and they remain in control of the data sharing process[1].

An integral data strategy should therefore be one of the building blocks of an ERN structure that aims to improve the life of people living with a rare disease through knowledge and data sharing across countries and diseases. **To fulfil this ambition, ERNs require strategic direction, a long-term vision and policy alignment with the wider European health data ecosystem.**

*Our aim is to trigger a strategic discussion on health data and ERNs that is long overdue.*

**This paper outlines EURORDIS' proposal to co-create a comprehensive ERN data strategy that matches the ambition of the Networks**, enabling them to address some of the most pressing public health needs of people living with a rare disease or complex condition.

## Why this paper?

Our aim is to trigger a strategic discussion on health data and ERNs that is long overdue.

**We propose to start the conversation by setting up a multi-stakeholder steering committee** that would address the following key questions:

1. What problems or questions do ERNs need to address that require the innovative use of health data?

2. What specific data is required to address the problems or questions that have been identified? Where is it stored? Is there a need to collect new data or can we reuse data sets already collected?

3. Are the data sets ready to be used at scale for data analysis and modelling? Do they meet the requirements in terms of quality, volume, availability and other relevant criteria?

4. Are the proposed uses of patients' data ethical and would they raise any regulatory concerns?

5. What data sharing scenarios (use cases) should be prioritised?

6. What are the technological and infrastructure requirements to support the prioritised scenarios?

In parallel, this multi-stakeholder steering committee should define an implementation framework fit for purpose. Our favoured scenario would see the Networks leading on the implementation strategy while relying on the support of a common infrastructure to streamline operations and build a common layer for project management, including a common health data governance framework, training for patients and clinicians, and changing management methodologies.

While we acknowledge that all 24 ERNs are different and might have different data needs, as well as different data readiness levels, we strongly believe that they will all benefit from a common implementation data infrastructure that can aggregate demand, provide project support and build common methodologies to support implementation, allowing for a certain degree of customisation when required.





We call on the ERN Board of Member States, the eHealth Network, the European Commission and the Network Coordinators to consider the proposals set out in this paper and take concrete steps to orchestrate a structured multi-stakeholder dialogue on this topic. It is time to recognise health data as a priority action for the ERNs and dedicate the resources and political push required to build a long-term vision that can effectively support the work of the Networks for the years to come.



# 1. Background and objective

Real progress on rare disease care and research is heavily dependent on our ability to pool and combine different types of data from various sources and across countries. **Health data**, coupled with **active patient organisations** and **connected clinical networks**, is the recipe to transform care delivery for rare diseases and drive research and innovation [2]. Today, the European Reference Networks (ERNs) bring together all these elements and this is why they represent a unique opportunity for the rare disease patient community.

ERNs are primarily a knowledge sharing and care coordination infrastructure, hence data sharing lies right at the heart of this new system that has been set up **to Share, to Care and to Cure**. The pace of progress for ERNs to fulfil their potential will depend on their ability to share data from different sources for multiple uses. This will remain the case regardless of whether they expand their activities to fully achieve each of the eight objectives reflected in the Cross-Border Healthcare Directive, or if they choose to focus their efforts on a number of core objectives[3].

*ERNs need a comprehensive data strategy to transform the wealth of health data and collective knowledge into improvements in the delivery of high-quality care. Lack of strategic direction in this critical area will lead to poor results in terms of scalability and long-term sustainability of the new system. Ultimately, it will result in a less than optimal contribution to the improvement of the quality of life of people living with a rare disease.*

**ERNs could be major contributors in shaping the future European health data ecosystem as they bring together three key ingredients to lead in this area:**

- Clinicians, patient advocates and patient organisations are already working together across EU borders towards a common goal: to improve care for rare diseases and advance research  
**(COMMON GOAL)**
- ERNs have started building-up the evidence base for the natural history of rare diseases, starting with the ERN registries and disease specific health outcomes. Rare disease patients are also willing to share their health data if the adequate safeguards and governance are in place  
**(DATA)**
- Clinicians and the rare disease patient community have a shared understanding of the importance of health data sharing and have developed a good degree of trust that is critical to articulate a sound health data governance framework for ethical and lawful health data sharing  
**(TRUST)**

However, this will not happen if decisions continue to be taken in an isolated way, on a case-by case basis, and investments fail to align with long-term strategic goals. If ERNs are in fact a new care structure, they **need to have a comprehensive health data strategy that is driven by a concerted policy action and is embedded in the wider European health data and IT ecosystem context in which ERN members operate.** The ultimate goal would be to have aggregated longitudinally patient level data that combines and links different types of data from different sources, for all rare diseases and across countries. While it is clear that this final objective cannot be sufficiently achieved by the Networks alone, they are in a privileged position to make meaningful progress towards this goal if they work in close collaboration with other stakeholders.

The conversation about the ERN data strategy must start **now**; if we continue delaying it, precious time will be lost with the risk that ad-hoc investment will continue to be made without a strategic direction and hence with a diminished impact.

Coordinating a data strategy for the ERNs will take time to build, not only because of the complexity of the topic and the environment, but also because we need to bring into the conversation a wide range of stakeholders and experts, beyond the clinicians and patient advocates involved in the Networks, the Board of Member States and European Commission.

In this regard, **EURORDIS encourages a joint approach where a vision for the ERNs data strategy is co-produced by rare disease experts alongside experts in digital health, and its implementation is funded by an adequate mix of project-funding and other funding instruments.** The objective of the proposals contained in this paper is to trigger a high-level strategic discussion on health data and suggest a course of action to build and deliver this strategic vision.





## 2. Why we need to act now

Rare disease patients want their data to support research and improve health outcomes.

The rare disease community is acutely aware of the importance of pooling data to advance research and improve health outcomes. **The vast majority of rare disease patients are willing to share their health data to foster research and improve healthcare, as long as they remain in control of their data throughout the data cycle and their choices and needs for information regarding the use of their data are respected** [1] (see detailed evidence in [Table 1](#)). Behind this readiness is the conviction that sharing data and combining different types of data from different sources will provide a multitude of insights and new knowledge that cannot possibly be derived from isolated databases.

We are aware of the challenges and the risks of extensive health data sharing and data linkage, but our community has expressed itself with a clear voice, and has also made clear the requirements and specific conditions that need to be in place to share their data (see [Table 1](#)). It is now a joint responsibility of policy makers, at EU and national levels, hospital managers, the scientific community, clinicians and patient organisations to respond to these needs and demands.

*The vast majority of rare disease patients are willing to share their health data to foster research and improve healthcare, as long as they remain in control of their data*

After a decade of policy development, an integral health data strategy for rare diseases is long overdue.

The 2008 European Commission communication on rare diseases[4] already highlighted the importance of databases and registries to increase knowledge on rare diseases and develop clinical research. Over the last 10 years, the European rare disease community has made substantial progress in defining rare disease data sets and data formats. However, despite a decade of policy developments and experts' discussions on rare disease data, we still lack an integrated approach to health data for the rare disease field.

A comprehensive rare disease data strategy is long overdue; a strategy that takes stock of the outcomes and lessons learned from [EPIRARE](#), [RD-CONNECT](#), the [EUCERD](#), the [Rare Diseases](#) Joint Actions and the numerous initiatives developed over the last 10 years, including [Orphanet](#) and the development of specific codification for the inclusion of rare diseases in national health information systems (OrphaCodes). Likewise, it should draw on the experience of the pilot projects developed by the partners of the Global commission to end the diagnostic odyssey of children living with a rare disease, as well as the work carried out by the World Economic Forum (WEF) initiative "Breaking barriers to health data"[5], [6].

Today, the structured cooperation enabled by the European Reference Networks makes them an excellent vehicle to address some of the unique challenges facing the rare disease community. Therefore, any future



rare disease data strategy should be implemented within this structure. It is time now, before we enter into the ERNs' next 5-year period, to transform that challenge into a priority action and take concrete steps to co-create a comprehensive data strategy that matches the ambitions of the Networks.

The degree to which this strategy will be successful will depend on our ability to co-create a common vision and on the availability of proportionate resources to deliver and implement it. Any future long-term financial planning for the ERNs should factor in the costs related to the development and implementation of this health data strategy.



## Policy fragmentation and siloed funding drags down the potential of the ERNs

Digital health policy and funding instruments remain fragmented in Europe, as the different European Commission Directorate Generals (DG CONNECT, DG SANTE, DG RTD and JRC) hold different competencies and goals in the area, and alignment with the national digital health strategies has not yet been fully achieved.

The policies of the three DGs respond to different logics, namely market growth and competitiveness (DG CONNECT), protecting and improving public health (DG SANTE) and scientific research (DG RTD). Also,

*We need greater  
coordination and more  
permanent collaboration*

their policies are funded through different instruments and this will continue to be the case under the 2021-2027 multiannual framework (European Social Fund Plus, Horizon Europe and its Missions, the new Digital Europe Programme, all set to contribute to fund ERN health data-related activities).

While the diversity of funding instruments is not a problem in itself, we believe that funding for infrastructure, such as ERN registries, should be articulated through a stable block funding mechanism that is not competitive. The implications of taking a competitive call approach, regardless of the type of

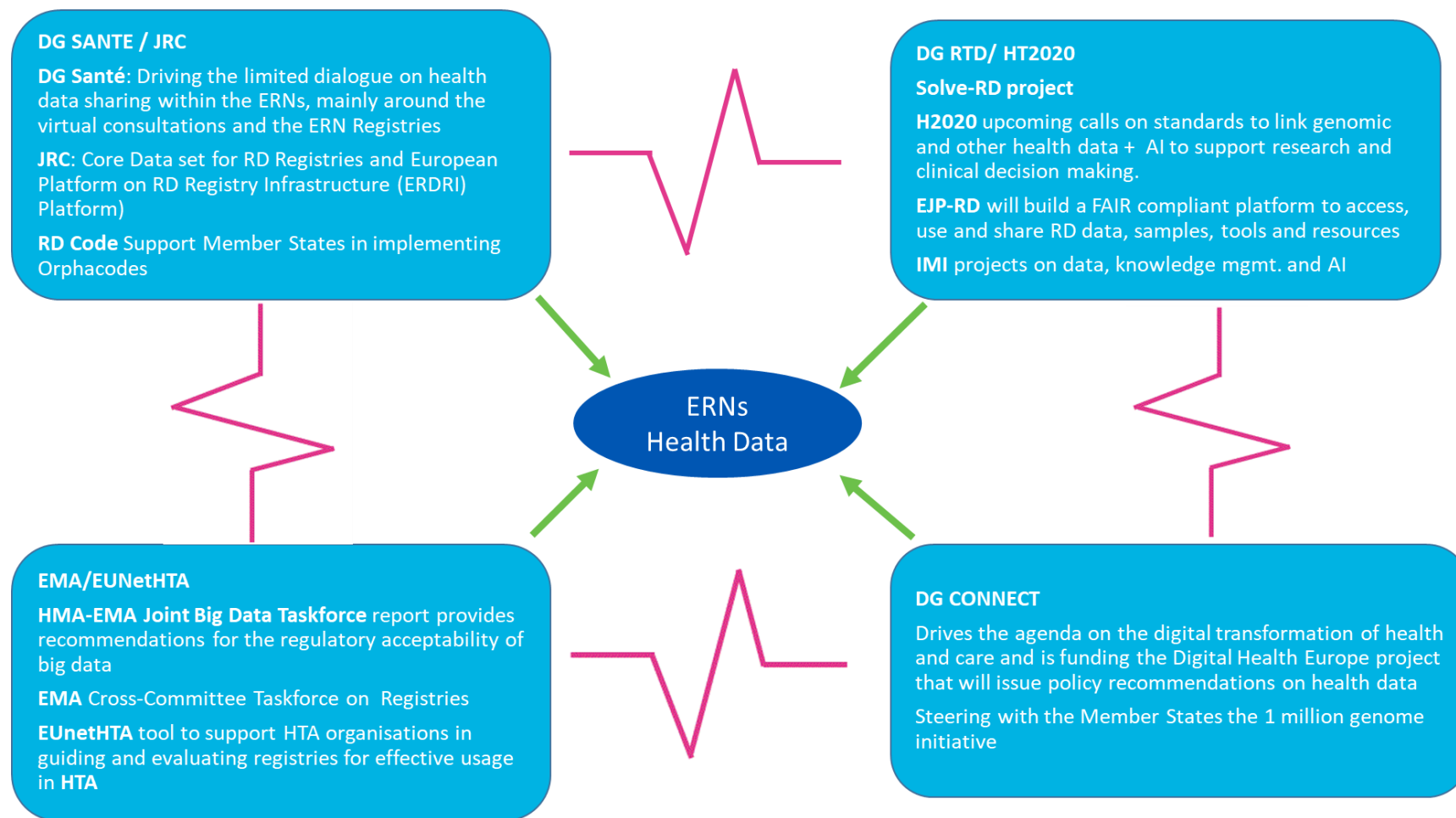
activity or service, automatically creates divergence and variability for no real benefit and increases the risk that ERNs capacity be solely focused on continually responding to new grant and funding applications.

Having different portfolios and a variety of funding instruments is positive as long as there is a close collaboration across portfolios. Creating a [temporary taskforce](#) for the development of health and digital policies was a first step in this direction, but we need greater coordination and more permanent collaboration structures between the involved Directorate Generals. Alternative instruments or solutions could be considered: for example, the R&I Missions' approach under Horizon Europe could be a step in the right direction as it would help break down silos, overcome funding fragmentation and favour cross-disciplinary and cross-sectoral innovation. At any rate, even though a "Mission" for a European health data ecosystem may not be the right tool, we certainly need concerted policy to orchestrate action in this field.

Overall, the different EU initiatives that involve the use of health data, of which ERNs should be a central piece, lack a comprehensive strategic vision. The different portfolios and funding instruments should not prevent the Commission from developing an integral data strategy for ERNs that is also connected with the wider policy agenda on the digital transformation of health and care.



Figure 1. Siloed approach to health data policy and funding affecting the European Reference Networks



**No time to lose: Building a data strategy for the European Reference Networks**

A EURORDIS contribution | April 2020

## Building a thriving ERN health data ecosystem requires a system approach and a long-term strategic vision. Yet, current experience goes in the opposite direction.

By way of example, the recent modification of the EC implementing decision on ERNs[7] includes new provisions that will now govern the secondary use of patients' data that has been entered into the Clinical Patient Management System (CPMS). When a patient consents to his or her data also being entered in *rare diseases registries or databases*, only healthcare professionals and *other individuals affiliated to the healthcare providers who are entitled to have access to patients' data* may use the data to identify similar patient cases.

Ideally, the wealth of data entered into the CPMS should be also made accessible and fully utilised for research purposes, including clinical trials and observational research[8]. However, the EC implementing decision now limits the secondary use of the data (it can only be used to *identify similar patient cases*) and raises other questions; it is unclear who the *other individuals affiliated to healthcare providers* are and it is also unclear what these *registries or other databases* are. It is also uncertain if these provisions would allow cross-ERN data sharing. It seems that we are about to create yet another data silo, one that only the healthcare professionals involved in a given ERN and "other individuals affiliated" to them will be able to access.

*The wealth of data entered into the CPMS should also be made accessible and fully utilised for research purposes*

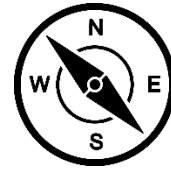


The ERN registries appear to be yet another example of how decisions on data have been taken on isolation and without a long-term vision. ERN registries are bound to be a central piece of the future ERN data strategy. However, not all ERNs have had sufficient time and resources to carefully consider aspects such as the long-term goals of their registry, its maintenance and evolution, links with other disease-specific registries, including patient-led registries, and how the ERN registry will fit into their overall data strategy.

All these decisions will affect the future rare disease health data landscape and should have been with the involvement of all those concerned, including patient organisations, some of which curate valuable data sets.



## Principles for a thriving ERN health data ecosystem



**Overseen by a patient-centric governance framework** that respects patients' preferences, allows them to take control of their data and involves rare disease patient organisations throughout the health data cycle.

**Based on a long-term strategy** that takes stock of previous initiatives and is co-created by a group of experts from the rare disease field alongside experts in digital health.

**Guided by the notions of data findability, accessibility, interoperability and reusability** (FAIR principles [9]) and supported by a patient identity management system to link individual patient data across multiple data sets.

**Enabled by an implementation framework** that combines common elements of support to all Networks with specific arrangements that can be set up by each Network.

**Driven by ERNs data needs** and at the same time **well anchored in the wider European health data ecosystem** and aligned with the goals of the European health data space and the work of the European Joint Programme on Rare Diseases (EJP RD).

**Steered by long-term operational approach** where data sets are curated and maintained over time and leverages on existing data patient-led and other health data platforms.

**Underpinned by an adequate incentive structure** that favours collaboration and recognises the value of data collection, data curation and data management.

**Adequately funded based on a realistic estimate of the costs** associated to collecting, curating and exploiting high quality data sets.

**Supported by an adequate mix of project-funding, grants and other type of funding instruments and in-kind contributions from different sources, including patient organisations and industry**, that provides stable resources for infrastructure and operations and allows for long-term financial planning.

### 3. What are the health data needs of the ERNs?

The last two years of health data sharing in the ERNs has focused primarily on the use of data to support the virtual consultations performed with the support of the CPMS. However, the potential for health data usage in the ERNs is considerably greater and spans beyond cross-border virtual consultations (see Figure 2). In order to start building an ERN health data strategy, the very first question that we need to address is what do ERNs want to do with the data or, in other words, what are the ERNs' data needs?

*The potential for health data usage in the ERNs is considerably greater and spans beyond cross-border virtual consultations*

Figure 2 below is not meant to represent an exhaustive taxonomy of health data types and data uses. It aims to show that clinical data is only a piece of the puzzle, that health data uses have no clear boundaries, and shows the value of aggregating different types of data from different sources to serve multiple needs (data uses). In most cases, data sets will be used in an aggregated and de-identified format. For the purposes of this paper, the term “de-identification” includes the full spectrum of methods, from simple pseudonymisation to full anonymisation (see definitions of these terms in Annex I).



Figure 2: Types of health data and data uses



Source: Own design based on the taxonomy of health data developed by Susannah Fox [10]

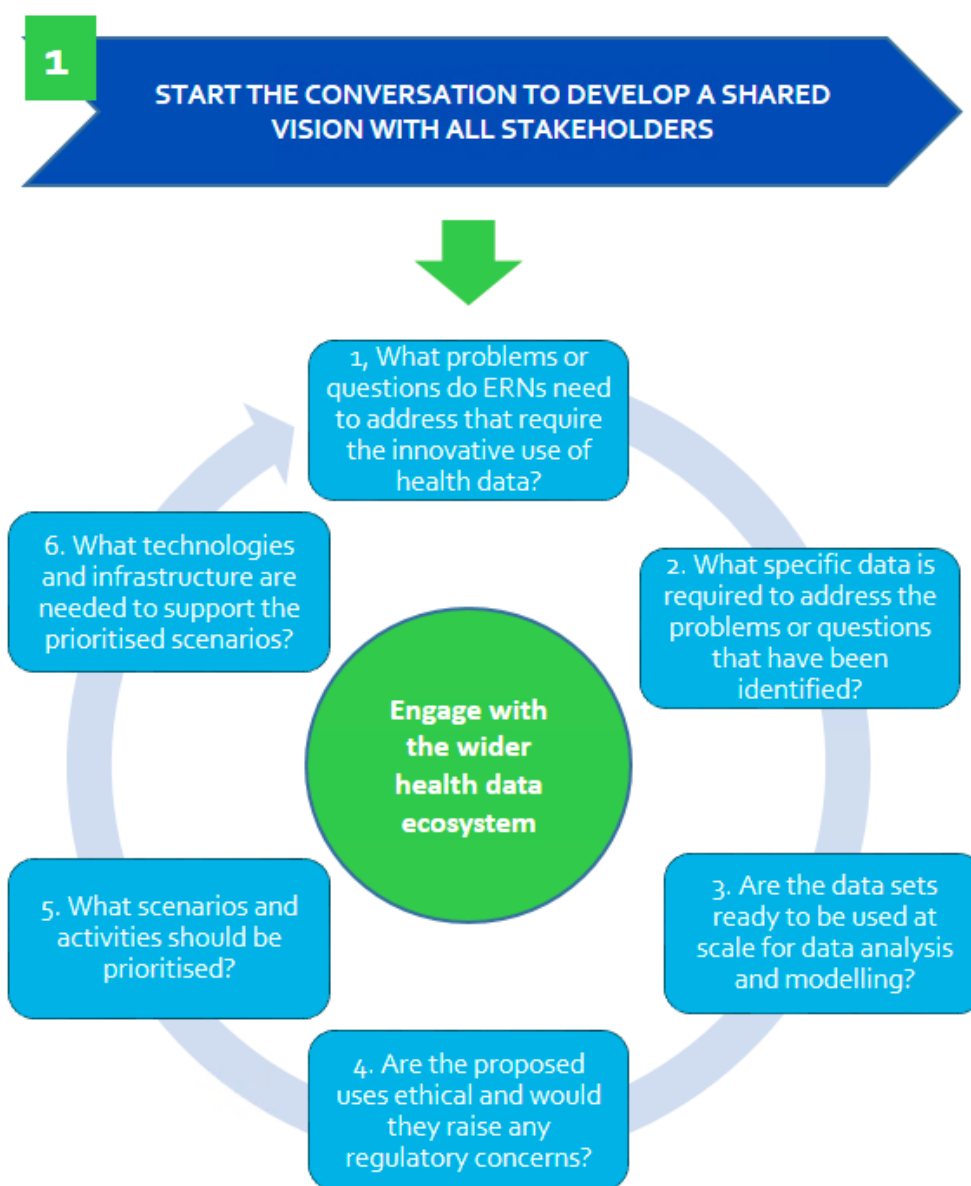
**No time to lose: Building a data strategy for the European Reference Networks**

A EURORDIS contribution | April 2020

## 4. Defining and delivering a health data strategy for the ERNs

EURORDIS would like to put forward a constructive proposal to progress towards developing a comprehensive ERN data strategy and a fit-for-purpose implementation framework. While the proposal is structured in two distinct sections, i.e. the strategic thinking and the operational aspects, both dimensions are closely interlinked and should be addressed at the same time.

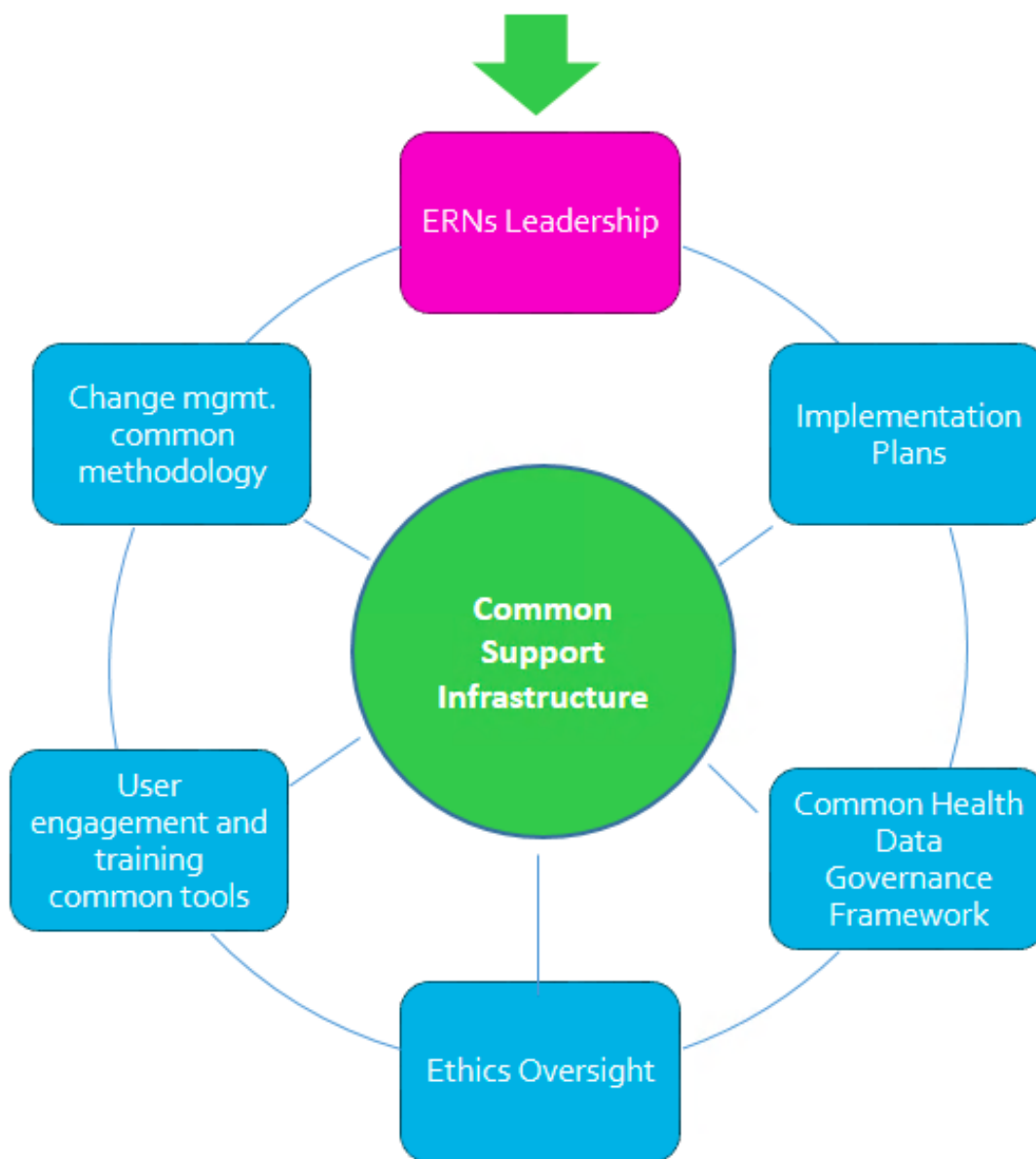
Figure 3: EURORDIS vision to progress towards an ERN Data Strategy





2

## SET UP A DELIVERY FRAMEWORK



# 1. Start the conversation to develop a shared vision with all stakeholders

ERNs do not exist in a vacuum; their members are healthcare providers established in 26 different Member States, who are subject to their own national legislations and specificities regarding the use of health data and who are rooted in their national, regional and hospital digital health strategies. Any decisions around health data should take into account the reality in which ERN members operate on a daily basis and seek alignment with external rare disease data infrastructures. In addition, there should be a genuine effort to articulate the mechanisms to link ERNs with the national and European digital health strategies. After all, some of the challenges linked to the cross-border exchange of health data will be the same or very similar regardless of the situation (emergency, routine or specialist care), as will the solutions.

ERNs need to develop a common understanding of their goals in terms of data usage, medium and short-term priorities and what is needed to deliver them. To achieve this and **to build a common vision that works for everyone, the ERN health data strategy would need to be developed by a multi-stakeholder steering committee with representation from the following groups:**

- Board of Member States representatives and other representatives from the Ministers of Health with responsibilities in digital health and data.
  - ERN members and patient representatives involved in the ERNs
  - Patient Organisations representatives
  - Hospital managers
  - EJP RD representatives
  - Orphanet
  - JRC representatives
  - European Commission representatives
  - Industry representatives
- External experts such as bio-informaticians, data scientists, IT and security experts, ethicists and data protection legal experts, etc.

**To start the conversation, we suggest that the multi-stakeholder steering committee addresses six initial questions:**

1. What problems or questions do ERNs need to address that require the innovative use of health data?

2. What specific data is required to address the problems or questions that have been identified? Where is it stored? Is there a need to collect new data or can we reuse data sets already collected?

3. Are the data sets ready to be used at scale for data analysis and modelling? Do they meet the requirements in terms of quality, volume, availability and other relevant criteria?

4. Are the proposed uses of patients' data ethical and would they raise any regulatory concerns?

5. What data sharing scenarios (use cases) should be prioritised?

6. What are the technological and infrastructure requirements to support the prioritised scenarios?

Addressing these questions in a structured way will provide good foundation to develop a shared vision, guiding principles, and agree on specific actions to move forward.

### **1. What problems or questions do ERNs need to address that require the innovative use of health data?**

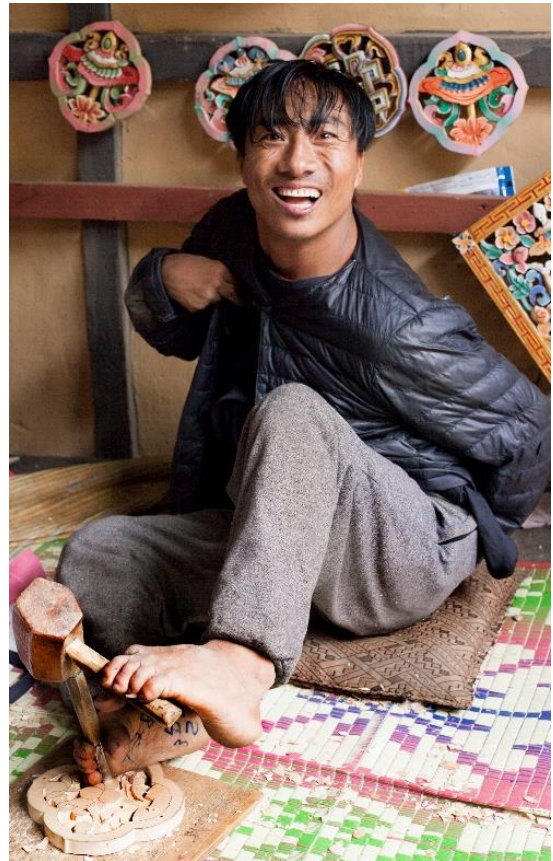
The discussion around potential uses should have enough level of granularity (down to the level of use cases) to be able to identify in the next stage what specific data is required. The objective at this stage will be to capture the views of the different stakeholders and agree on a common vision around the purposes that works for everyone. While there will be probably be a general agreement on the overarching purpose (using data to improve health, care and services), the specific priorities of each stakeholder may differ. Questions around the incentive structure for data sharing should also be explored at this point to address concerns about access to the data and eventually to the results of the data analysis.

## 2. What specific data is required to address the problems or questions that have been identified? Where is it stored? Is there a need to collect new data or can we re-use data already collected?

As part of this conversation, we should reflect on what type of data is required to develop the services and tools identified in step 1.

In addition, there will be a need to map existing data sets and understand the rules governing access and use in each case. This will also allow to decide whether it is possible or not to combine different data sets. For example, the data from the UK 100,000 genomes project cannot be shared. It can be studied, but research must take place within [Genomics England's](#) secure servers, with only results and analysis being withdrawable. These trusted research environments might become more common-place as a way to manage consent and governance of use of large sets of identifiable data. The immediate question that would need to be addressed is what happens when a research project wants to combine two data sets, when neither is allowed out of its secure environment.

Finally, at this stage, early consideration of how to incentivise data collection and ensure data quality will help to address concerns and questions about extra burden of work.



## 3. Are the data sets ready to be used at scale for data analysis and modelling? Do they meet the requirements in terms of quality, volume, availability and other relevant criteria?

To answer this question, the steering committee members will first need to understand the characteristics of the data sets required to address the needs identified in step 1.

To obtain value from the data there is **a substantial amount of work that needs to be done to prepare the data sets for modelling and analysis**. Before any analytics and data linkage can happen at scale, the raw data needs to be curated, cleansed, mined and normalised. The goal would be to have longitudinally patient level data that combines and links data from different data sets (EHR and registries for example), different types of data (phenotypic and -omics data for example) and data sets from different care settings (primary, secondary, and tertiary health and medical records).

At this stage, the steering committee should at least be able to look into the following four dimensions to determine the characteristics of the data sets, whether they meet the criteria to be used for the use cases identified in step 1 and to what extent they comply with FAIR data principles:



## 1 Nature

- **Data type** (patient reported, clinicians/researchers, hybrid -captured by patient and validated by professionals)
- **Data availability or time frame** (RWE vs. historical observational data with time lag)
- **Source** (available from a single source vs. multiple)
- **Granularity or detail** (aggregated vs. transaction level)

## 2 Data quality, maturity and embedded with analytic insight

- Raw (unorganised with potential data gaps and inconsistencies)
- Curated (i.e. organised and easy to work with)
- De-identified and level of de-identification
- Aggregated longitudinally for the same patient or record)
- Compliant with FAIR data principles[9]
- Analysed with descriptive statistics, insights and predictions and forecast provided

## 3 Complexity of data capture

- Accessibility of the data (paid vs. open source)
- Data capture (auto captured vs. collected with human intervention)
- Feasibility of using algorithms to process, convert and aggregate data already collected and from multiple sources.

## 4 Use/application

- Use and potential impact
- Limitations on use
  - Exclusivity (exclusive licence vs. data offered to multiple actors)

Source: Adapted from the Framework for analysing characteristics that impact the value of a data set[11]



#### 4. Are the proposed uses ethical and would they raise any regulatory concerns?

At this stage, the steering committee should determine whether the selected uses respect fundamental rights and common values. Prospectively evaluating the possible effects of the selected uses on patients, society and the common good. In addition, the legal and regulatory requirements should be examined at this stage – how can we use the data for the purposes identified in step 1 while remaining compliant with the European and national data protection laws? Are there any health data anonymisation standards that should be observed? Are the legal requirements such, that some of the identified uses are not viable?

#### 5. What data sharing scenarios (use cases) should be prioritised?

Based on the readiness of the data and the analysis of the ethical and regulatory concerns, the group will be able to prioritise some scenarios (*use cases*), for example accurately extracting phenotypes from heterogeneous EHR notes, and defining the actions and tasks required to exploit the data.

Adopting an incremental approach by selecting a number of scenarios to show how collaboration can work will already be good progress. The guiding principle could be to aggregate demand and address in the first place the scenarios that will address the needs of all ERNs, or a substantial number of them. There appears to be enough common challenges to be able to identify actions that will be relevant for all ERNs and will help to get them to the same level playing field in terms of data collection, data governance, implementation of OrphaCodes, etc. In addition, if more than one ERN is interested in using the same data set, there will be an efficiency in centralising actions. This will be a strong benefit of the strategy.

We anticipate that preparing existing data sets (cleaning and normalising) and upgrading their quality will be one of the first major steps and progress towards this will already be a major achievement if it is done in the framework of a wider strategic planning. In addition, activities around [quality assurance of rare](#)

[disease registries](#), along the lines developed by the ENDO ERN registry project, might be required across the board.

## 6. What are the technology and infrastructure requirements to support the prioritised scenarios?

Among others, questions around the technology and infrastructure needed to support the delivery of the data strategy will include:

- Data acquisition/collection and cleansing (capturing, filtering and cleaning data);
- Data architecture: platforms/cloud-based solutions, availability, standards, privacy by design and security, performance and scalability, data models, data linkage, semantic interoperability, data extraction, etc.; Data curation: systems to support active management of the data, including data quality, provenance, etc.;
- Data usage: clinical decision support tools, knowledge sharing platforms, visualisation, communication standards;
- Patients and professionals' identity management.

## 2 Set up an implementation framework

Careful planning and agreement around the delivery methods is just as important as developing a shared vision and should be a key element of the ERN health data strategy. Various options could be envisaged here, ranging from a centralised approach (implementation strategy managed end-to-end by the European Commission in the same way as the CPMS project) to a decentralised approach, where each ERN defines its own implementation strategy, as it has happened with the ERN Registries. However, our preferred option envisages having a combined bottom-up, middle-out approach.

---

ERNs would lead on the implementation strategy (bottom-up), but would also benefit from a common support infrastructure (middle-out) to streamline operations and build a common layer for project management to avoid duplicating efforts, aggregate demand and allow to re-use assets whenever possible. This centralised support could help to define roles and responsibilities, estimate resources, provide support in the preparation of project proposals, develop common user engagement processes and change management methodologies. It would also provide a common health data governance framework, including ethics oversight. The work developed in the framework of the EJP RD for the virtual platform on data and resources could provide a good basis for this framework.

---

Certainly, a common health data governance framework will have to co-exist with the health data governance rules of each of the HCPs and those of other external infrastructures/platforms and databases. Nonetheless, there is still a need for an additional layer of governance that would apply when data is shared across HCPs and with other organisations in the activities performed under the umbrella of the ERNs structure. Even if different rules will apply depending on the specific use/service (who can access

the data, under what conditions, for what purposes, etc.), we believe that ERNs should have a common framework to help oversee and track the use of health data across the different ERN-related projects and services, to better manage the information and quality of the data.

**Such a framework should include at a minimum the following three dimensions:**

- 1)** Data stewardship, specifying the roles and responsibilities around data management and accountability.
- 2)** Data policies and procedures to manage data sets, including enforcing authentication and access rights to data as well as the organisational measures and policies to ensure the quality, accuracy and security of the data and regulatory compliance. Tools to help preserve the autonomy and rights of individuals to control their data.
- 3)** Data standards, specifications and rules for the definition, creation, storage and usage of data

The results from a recent Rare Barometer survey on health data sharing and data protection [1], could be used to inform this framework (see [Table 1](#)):



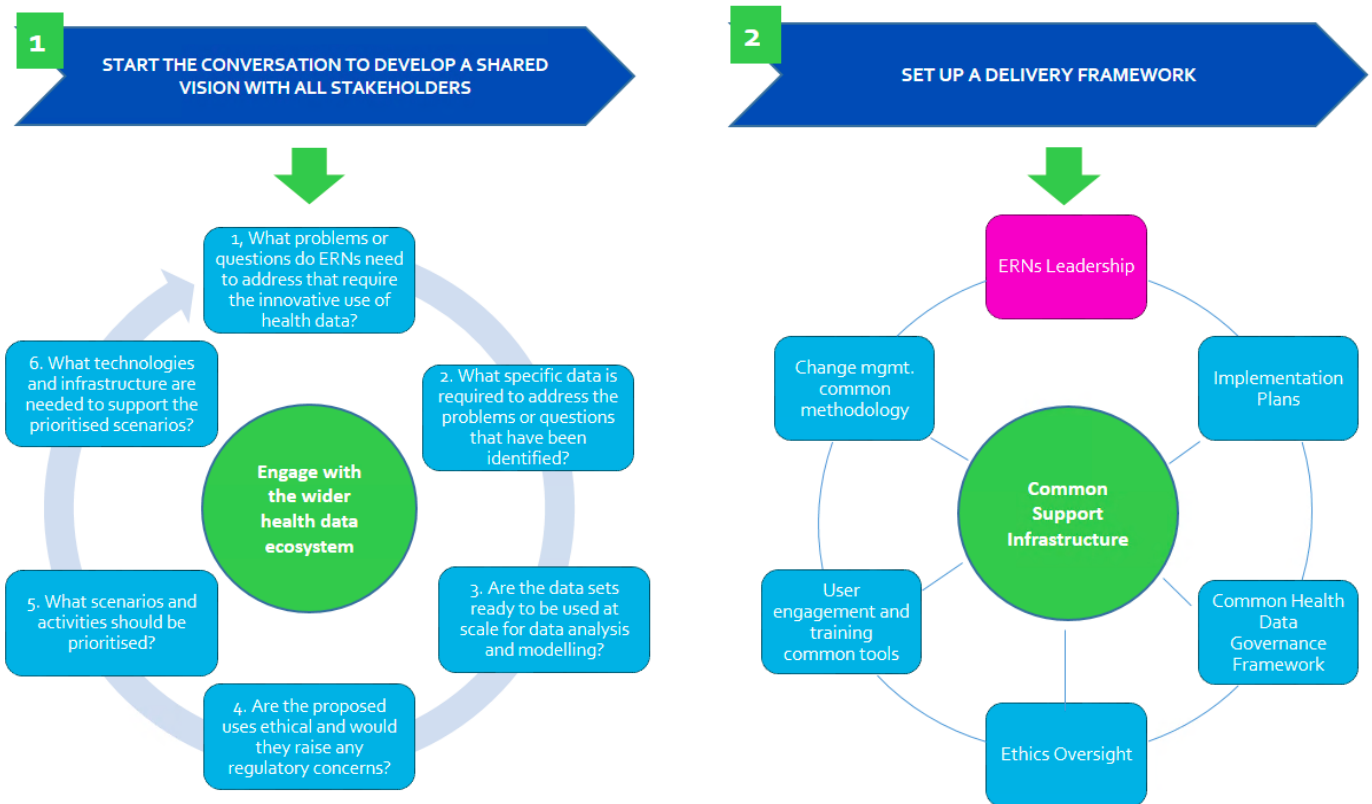
Table 1: Rare Barometer Survey on data sharing results and recommendations for health data governance (results based on 2013 responses)

Survey Results[1]	Principles for an ERN health data governance framework
Patients want to keep control of the data they are sharing (80% of respondents are in favour of keeping a high level of control)	1) <b>Move beyond obtaining valid consent to keep patients in control of their data</b> and explore how patients can be meaningfully involved across the data cycle, providing them the tools to understand the risks, express their preferences and act upon their data in an adaptive way. This idea is also linked to the role that Patient Organisations could play as data curators.
Trust in not-for-profit stakeholders to handle and use their health info (89% for medical doctors, 79% researchers from non-profit organisations, 77% for patient organisations, 69% for healthcare professionals other than medical doctors) is considerably higher than trust in for profit stakeholders.	2) <b>Meaningful involvement of medical doctors, researchers from non-profit organisations and patient organisations across the data sharing cycle</b> should be one of the <b>central guiding principles</b> underpinning the future ERN health data governance framework to ensure a shared responsibility in data stewardship and oversight. 3) <b>Equal partnership in decision-making structures</b> of clinicians, researchers and patient organisations. 4) <b>Patient Organisations and patients</b> contributing their data to a given platform should have <b>access to the curated</b> data on equal grounds and under the same conditions as any <b>other stakeholder</b> . 5) <b>Communication with patient organisations and clinicians</b> around all data strategy aspects (clear policies, standards, progress and results) should be built into the governance framework.
The uses of data under unchosen circumstances are the main risks associated with sharing data. 50% of respondents are concerned that their data could be used by third parties with which they would not have chosen to share their data.	6) <b>Ethics considerations</b> should be examined throughout the whole data cycle from inception (what data, for what purposes, risks and benefits, etc.) through implementation and follow-up (is the data achieving the transformation we had envisaged? are we seeing any impact on inequalities?) 7) <b>Address privacy and security needs</b> to ensure that only relevant data is shared in a safe and secure way and that there is a redress mechanism in place in case of misuse
90% of respondents are willing to share their data for wider scientific interest, to advance care and research not directly linked to their own disease	8) <b>Avoid data silos and exclusivity of use.</b> Patients want their data to be shared in a secure way, but at the same time, they want it to be shared widely to advance scientific research and care for rare disease patients.

## 5. Conclusion

ERNs need a comprehensive data strategy to transform the wealth of health data and collective knowledge into improvements in the delivery of care and a thriving rare disease research landscape. Lack of strategic direction in this critical area will lead to poor results in terms of scalability and long-term sustainability of the ERNs and will prevent the Networks from successfully delivering on their mission. The exchange of health data to support the virtual consultations is only the tip of the iceberg. The data needs of the ERNs will be greater and they should be clearly defined as a first step towards an integral ERN health data strategy.

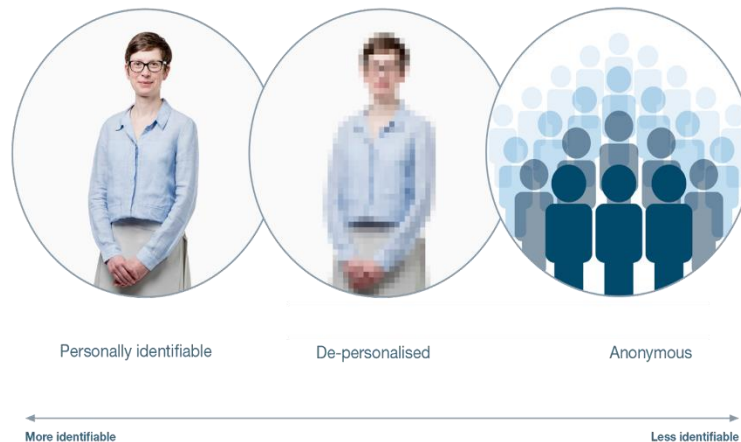
Our proposal to progress towards developing a comprehensive ERN data strategy is structured around two distinct dimensions, the strategic thinking and the operational aspects, both dimensions are closely interlinked and should be addressed at the same time.



Surely, we will not have the answer to all questions from the outset, but by delaying further the discussion around this critical topic, the gap between what we would like the ERNs to deliver and what they are actually delivering will just grow bigger, diminishing the opportunities that this new structure can offer us.

# Annex I. Definitions and De-Identification Terms

## Spectrum of Identifiability



Source: 'Identifiability spectrum' by [Understanding Patient Data](#), licensed under [CC BY](#).

## Personal Data

Any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data.

The General Data Protection Regulation (GDPR) protects personal data regardless of the technology used for processing that data – it's technology neutral and applies to both automated and manual processing, provided the data is organised in accordance with pre-defined criteria (for example alphabetical order). It also doesn't matter how the data is stored – in an IT system, through video surveillance, or on paper; in all cases, personal data is subject to the protection requirements set out in the GDPR [12].

## Patient Data

Data that is collected about a patient whenever they go to a doctor or receive social care. It may include details about the individual's physical or mental health, such as height and weight or detail of any allergies, and their social care needs and services received. It may also include next of kin information. This is recorded and stored in a care record [13].

## Anonymisation

The process of rendering data into a form which does not identify individuals either directly or indirectly and where identification is not likely to take place by any means reasonably likely [13].

There is a lot of research currently underway in the area of anonymisation, and knowledge about the effectiveness of various anonymisation techniques is constantly changing [14].

## Pseudonymisation

The process of replacing any identifying characteristics of data with a pseudonym, or, in other words, a value which does not allow the data subject to be directly identified. The GDPR defines pseudonymisation as the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that (a) such additional information is kept separately, and (b) it is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable individual.

Pseudonymisation should be distinguished from anonymisation, as it only provides a limited

protection for the identity of data subjects in many cases as it still allows identification using indirect means. Where a pseudonym is used, it is often possible to identify the data subject by analysing the underlying or related data.

Uses of anonymisation and pseudonymisation Data which has been irreversibly anonymised ceases to be “personal data”, and processing of such data does not require compliance with the Data Protection Law. In some cases, it is not possible to effectively anonymise data, either because of the nature or context of the data, or because of the use for which the data is collected and retained [14].

*Note: Additional terms and definitions around patient data can be found in this [Glossary](#) developed by Connected Health Cities.*



# Bibliography

- [1] S. Courbier, R. Dimond, and V. Bros-Facer, “Share and protect our health data: an evidence based approach to rare disease patients’ perspectives on data sharing and data protection - quantitative survey and recommendations,” *Orphanet J. Rare Dis.*, vol. 14, no. 1, p. 175, 2019.
- [2] “Bridging Patients and Researchers to Build the Future Agenda for Rare Disease Research in Europe. Major Workshop Outcomes,” in *Workshop co-organised by EURORDIS with e-RARE in partnership with European Commission, ORPHANET and EUROPLAN*, 2010.
- [3] *Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients’ rights in cross-border healthcare. Art.12. .*
- [4] *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Rare Diseases: Europe’s challenges. 11.11.2008 COM(2008) 679 final. .*
- [5] World Economic Forum, “Breaking Barriers to Health Data Governance Workshop - Workshop Report - April 2019-min.”
- [6] World Economic Forum, “Quarterly Project Update: Breaking Barriers to Health Data,” 2019.
- [7] *Commission Implementing Decision (EU) 2019/1269 of 26 July 2019 amending Implementing Decision 2014/287/EU. .*
- [8] EU Joint Action on Rare Cancers, “Rare Cancer Agenda 2030: Ten Recommendations from the EU Joint Action on Rare Cancers,” 2019.
- [9] M. D. Wilkinson *et al.*, “Comment: The FAIR Guiding Principles for scientific data management and stewardship,” *Sci. Data*, vol. 3, Mar. 2016.
- [10] “Mapping the health data landscape.” [Online]. Available: <https://susannahfox.com/2019/06/17/mapping-the-health-data-landscape/>. [Accessed: 14-Nov-2019].
- [11] EY, “Realising the value of health care data: a framework for the future.” 2019.
- [12] European Commission, “What is Personal Data?” [Online]. Available: [https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data\\_en](https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en). [Accessed: 19-Nov-2019].
- [13] “Data Glossary - Connected Health Cities.” [Online]. Available: <https://www.connectedhealthcities.org/community/glossary-data-use/>. [Accessed: 14-Nov-2019].
- [14] Irish Data Protection Commission, “Guidance Note : Guidance on Anonymisation and Pseudonymisation,” no. June. 2019.

# Additional References

Agency for Health Care Program (2012). “Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition”. Retrieved from <https://effectivehealthcare.ahrq.gov/products/registries-guide-3rd-edition/overview/>

Chediak, L. (2019). “World Economic Forum Quarterly Project Update: Breaking Barriers to Health Data”. Webinar, retrieved from <https://www.weforum.org/>

Commission Expert Group on Rare Disease (2014). “[Recommendation on Ways to Improve Codification for Rare Diseases in Health Information Systems](https://ec.europa.eu/health/sites/health/files/rare_diseases/docs/recommendation_coding_cegrd_en.pdf)”, adopted by the Commission Expert Group on Rare Diseases”. Retrieved from [https://ec.europa.eu/health/sites/health/files/rare\\_diseases/docs/recommendation\\_coding\\_cegrd\\_en.pdf](https://ec.europa.eu/health/sites/health/files/rare_diseases/docs/recommendation_coding_cegrd_en.pdf)

eHealth Network (2019). “eHealth Network Guidelines on an interoperable eco-system for digital health and investment programmes for a new/updated generation of digital infrastructure in Europe”. Retrieved from [https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev\\_20190611\\_co922\\_en.pdf](https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20190611_co922_en.pdf)

Elixir (2019). Retrieved from <https://elixir-europe.org/>

EUCERD (2013). “EUCERD Core Recommendations on Rare Disease Patient Registration and Data Collection”. Retrieved from [http://www.eucerd.eu/wp-content/uploads/2013/06/EUCERD\\_Recommendations\\_RDRegistryDataCollection\\_adopted.pdf](http://www.eucerd.eu/wp-content/uploads/2013/06/EUCERD_Recommendations_RDRegistryDataCollection_adopted.pdf)

EU Joint Action RD-ACTION (2012-2015). Task-Force on Interoperable data-sharing in Rare Diseases and eHealth communities. Retrieved from <http://www.rd-action.eu/ehealth-and-european-reference-networks/>

[EU Joint Research Centre](http://www.erare.eu/sites/default/files/SetCommonData-EU%20RD%20Platform_CDS%20_final.pdf) (2017) “set of common elements for RD registration”. Retrieved from [http://www.erare.eu/sites/default/files/SetCommonData-EU%20RD%20Platform\\_CDS%20\\_final.pdf](http://www.erare.eu/sites/default/files/SetCommonData-EU%20RD%20Platform_CDS%20_final.pdf)

European Commission (2019). “Commission Recommendation on a European Electronic Health Record exchange format (C(2019)800)”. Brussels

European Data Protection Supervisor, “A Preliminary Opinion on data protection and scientific research,” 2020.

European Medicines Agency (launched 2015) “Initiative for patient registries”. Retrieved from <https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries>

EURORDIS (2013). “[The Voice of Rare Disease Patients: Experiences and Expectations of over 3,000 Patients on Rare Disease Patient Registries in Europe](https://www.eurordis.org/publication/voice-rare-disease-patients-experiences-and-expectations-over-3000-patients-rare-disease-patient-registries-europe)”. Retrieved from <https://www.eurordis.org/publication/voice-rare-disease-patients-experiences-and-expectations-over-3000-patients-rare-disease-patient-registries-europe>

Heads of Medicines Agencies (HMA) and European Medicines Agency (EMA) Joint Task Force to establish roadmap and recommendations for use of big data in assessment of medicines (launched 2017). Retrieved from: <https://www.ema.europa.eu/en/about-us/how-we-work/big-data>

OECD (2015). “Health Data Governance – Privacy, Monitoring and Research”. OECD Health Policy Studies. Retrieved from <https://www.oecd.org/publications/health-data-governance-9789264244566-en.htm>

OECD (2019), Health in the 21st Century: Putting Data to Work for Stronger Health Systems, OECD Health Policy Studies, OECD Publishing, Paris, <https://doi.org/10.1787/e3b23f8e-en>.

OECD (2019). Health in the 21<sup>st</sup> Century. Retrieved from <http://www.oecd.org/health/health-in-the-21st-century-e3b23f8e-en.htm>

Phillips M., Knoppers B.M., Baker D. & Kaufmann P. (2018). “Privacy-Preserving Record Linkage: Ethico-Legal Considerations”. IRDiRC, retrieved from <http://www.irdirc.org/activities/task-forces/privacy-preserving-record-linkage/>

Understanding Patient Data (2019). “What are the best words to use when talking about data?”. Retrieved from <https://understandingpatientdata.org.uk/what-are-best-words-use-when-talking-about-data>



#### **EURORDIS-RARE DISEASES EUROPE**

Plateforme Maladies Rares ♦ 96 rue Didot  
75014 Paris ♦ France

#### **EURORDIS Brussels Office**

Fondation Universitaire ♦ Rue d'Egmont 11  
1000 Brussels ♦ Belgium

#### **EURORDIS Barcelona Office**

Recinte Modernista Sant Pau ♦ Pabellón de Santa Apolonia  
Calle Sant Antoni M<sup>a</sup> Claret 167 ♦ 08041 Barcelona ♦ Spain

**EURORDIS.ORG**

This paper was produced thanks to funding received under an operating grant from the European Union's Health Programme (2014-2020).



Co-funded by  
the Health Programme  
of the European Union

