# HealthData@EU Pilot identifies common elements for health data access and data use within the legal frameworks of the participating nodes

**Launched in October 2022, the HealthData@EU Pilot project has recently completed a landscape analysis of legal frameworks for health data collection and health data use underpinning the five cross-border research use cases of the project. The document was prepared by Work Package 7 "regulatory and legal compliance", led by BBMRI-ERIC and provides an overview of the main legal and regulatory similarities and differences observed within the network.**

The landscape analysis aims to identify common elements and major differences between the national health data access bodies, Research infrastructures and EU agencies involved in the project which are legally and technically competent to gather and provide health data (nodes) and gather documentation (e.g. existing data application forms, data use policies, etc.).

The analysis will be used as groundwork for the next activities planned within Work Package 7, including the design of a common data application form as well as aligned conditions for data use.

A detailed questionnaire was filled in by the following nodes: BBMRI (European Research Infrastructure), Health Data Lab (Germany), Danish Health Data Authority (Denmark), Findata (Finland), Health Data Hub (France), Sciensano (Belgium), Norwegian Directorate of eHealth(Norway) and Croatian Institute of Public Health (Croatia).

## Main takeaways

### Overall conditions and scope

- Nodes vary to a great extent regarding the scope of available data. All nodes seem to use exclusively their already existing data collections (either built for the administration of the respective health system or, in the case of research infrastructure, based on existing research databases). No node has established or so far considered a general collection of health data to make it comprehensively available to policy or research.
- Most of them have a focus on data stemming from the health care system.
- There is no node that has data from industry, neither from clinical trials nor from medical devices.
- Some nodes do not only make health data available, but also data from other social security systems such as unemployment or pension data.
- The statutory purposes for which the health data may be processed in the nodes correspond largely to the purposes listed in Art. 33 of the draft EHDS regulation.
- In most nodes, there are multiple legal bases for making health data available.

### Data discovery

- An exhaustive metadata catalogue seems to be rather the exception than the rule but most nodes are undertaking efforts to build or consolidate a metadata catalogue.

Co-funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

**Data access application and permit**

- All nodes authorise the following categories of actors to access data: legally defined bodies with a policy function (eg. regulatory tasks or disease surveillance) and authorised researchers.
- The possibility to apply in the English language is rather the exception than the rule.
- In all cases there is a digital application platform and most nodes provide for time limits in handling applications which can take from one to six months in practice.
- In most nodes an ethical/scientific committee is involved. However, their role is quite heterogeneous in the different Member States. Data protection authorities are usually not involved in the delivery of the permit, with one notable exception in the French case.

**Contractualisation**

- Rules about fees and credits vary to a great extent.

**Data preparation, provision and use**

- All countries facilitate data linkage, for instance to unemployment or pension data using probabilistic methods or a national unique identifier, and in collaboration with other national administrations.
- Health data is generally made available in pseudonymised format, often in a secure processing environment.
- Timelines for data provision are long as often no strict time limits exist. Depending on the complexity of the request, the provision of data can take up to one year after the permit has been delivered.

**Citizen engagement and citizen GDPR rights** are implemented very heterogeneously across the nodes.

- For instance, some nodes build on informed consent while others propose an opt-out system. The same is true for transparency. Measures range from individual information, to collective information via website to no information at all.
- Many measures to exercise GDPR rights are still being implemented or even debated.

Overall, it can be said despite some commonalities, the user journey for accessing health data in the nodes involved differs in many ways. The differences range from the type of data available, and from the legal basis for making it available to the process of making it available and the response to data misuse. There are also major differences in the degree of centralisation of the decision on data access. The systems considered range from largely centralised access decisions to largely decentralised systems.

## What are the next steps?

These findings, as well as the documentation gathered (e.g. existing data application forms, documentation to help the application, metadata catalogues, architecture of existing portals, data use/sharing policies) will allow identifying good practices, similarities and major pitfalls associated with each stage of health data discovery, health data access application process and use of health data, and will feed into the next outputs of WP7:

- Design of a common data application form leaving some leeway for nodes differences;
- Definition of general conditions of data use and security measures;
- Development of a model for data use agreement between nodes and data users.

# Overall conditions and scope

## Scope of data available

Nodes vary to a great extent regarding the scope of available data.

- All nodes seem to use exclusively their already existing data collections (either built for the administration of the respective health system or - in the case of research infrastructures - based on existing research databases).
- Most of them have a focus on data stemming from the health care system. For example, these nodes include health registry data or social insurance data that constitute the core of the medical databases and data from electronic patient records.
- Another more or less important data source are disease registries such as cancer registries and public registries such as vaccination or mortality registries.
- Finally, some nodes do not only make health data available, but also data from other social security systems such as unemployment or pension data.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Data stemming from the health care context | | ✓ | ✓ | ✓ | ✓ | | | ✓ |
| Public health registries | | | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Data stemming from research context | ✓ | | | ✓ | ✓ | | ✓ | |
| Other categories | | | | ✓ | ✓ | | | ✓ |

## Legal basis for making health data available

In most nodes, there are multiple legal bases for making health data available (either per categories of data, or per purposes).

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| One overall legal allowance comparable to EHDS Regulation | | | | ✓ | | | | |
| Multiple legal references per data source and/or purpose | | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| No special legal allowance, therefore consent of data subject | ✓ | | | | | | | |

# Data discovery

An exhaustive metadata catalogue exists for BBMRI, Finland and Norway, but it seems to be rather the exception than the rule.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| A metadata catalogue exists | ✓ | | | ✓ | | | ✓ | |
| A metadata catalogue partially exists | | | | | ✓ | ✓ | | |
| There is no metadata catalogue | | ✗ | ✗ | | | | | dev. |

*In the interest of readability, the following pictogrammes will be used:*

| | BBMRI | HDL | DHDA | Findata | HDH | Sciensano | NDEH | CIPH |
|---|---|---|---|---|---|---|---|---|
| Logo | BBMRI-ERIC | | DANISH HEALTH DATA AUTHORITY | FINDATA | HEALTH-DATA-HUB | sciensano | | CROATIAN INSTITUTE OF PUBLIC HEALTH |
| Pictogramme | | | | | | | | |

# Data permit / access application

## Data users

All nodes authorise the following categories of actors to access data: legally defined bodies with a policy function (eg. regulatory tasks or disease surveillance) and authorised researchers.

Some nodes only grant access to applicants from abroad if they collaborate with national institutions.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Everybody serving the legal purposes | | | | ✓ | ✓ | | ✓ | ✓ |
| Legally defined policy makers | | ✓ | ✓ | | ✓ | ✓ | | |
| Identified and authorised researchers | ✓ | ✓ | ✓ | | ✓ | ✓ | | |
| Industry as researcher | ✓ | | | | ✓ | ✓ | | |
| Industry only in collaboration with academic researchers | | | ✓ | | | | | |
| Global researcher | ✓ | | | | ✓ | ✓ | | |

## Data access application

The possibility to apply in the English language is rather the exception than the rule.

In most cases there is a digital application platform, where the following documents have to be submitted:

- application form
- proof of affiliation
- description of requested data
- study protocol/purpose description
- ethics vote or other approvals/permissions.

Finally, most nodes provide for time limits in handling applications.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Application in English possible | ✓ | | | ✓ | | ✓ | ✓ | ✓ |
| Request portal up and running | ✓ | in dev | ✓ | ✓ | ✓ | ✓ | ✓ | |

## Data permit

In most nodes an ethical/scientific committee is involved.

For instance, the French ethical and scientific committee checks that the purpose of the study is relevant and of "public interest", that the data requested is appropriate and that the proposed methodology is robust, and supplies a non-binding opinion. In Norway, the Regional Ethics Committee for Medical and Health Research (REK) assesses the ethics in protocol/projects if defined as medical or health research, and the competency and formalities of involved coworkers and responsible parties (project leader, data protection, responsible parties etc.). The process of ethical approval takes 2+ months.

In Denmark, a Danish National Centre for Ethics has been set up in 2021 and supports the work of four independent bodies (Danish Council on Ethics, Danish National Committee on Health Research Ethics and Danish Medical Research Ethics Committee).

Data protection authorities are usually not involved in the delivery of the permit, with one notable exception in the French case.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Review by an ethical/scientific committee, DPA or similar through data permit authority | ✓ (optional) | | | | ✓ | ✓ | | ✓ |
| Ethics vote of another ethics committee acknowledged | ✓ | | ✓ | ✓ | | | ✓ | |
| Involvement of the DPA in the process | | | | if needed | Grants permit | x | | Review |

# Contractualisation

## Fees and credits

Rules about fees and credits vary to a great extent. For instance, in Denmark, France and Belgium, there is no fee for public authorities, whereas in Germany, Finland and Croatia, there are fees for all applications.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| No fees at node level | ✓ | | | | | | | |
| Depends on actual data controller | ✓ | | | | ✓ | ✓ | ✓ | |
| No fees for public authorities, but for researchers for the administrative effort | | | ✓ | | ✓ | ✓ | | |
| Fees for all applicants | | ✓ | | ✓ | | | ✓ | ✓ |
| Data providers are reimbursed | | | | ✓ | | | | |

# Data preparation, provision and use

## Data linkage

Smaller countries, such as Denmark, Finland, Belgium and Croatia are more likely to have a national unique identifier, whereas France and Germany seem to have reservations from the data protection perspective against such a number.

But all countries facilitate data linkage for example to unemployment or pension data using probabilistic methods if needed, and in collaboration with other Nodes.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Nationwide identifier exists | | | ✓ | ✓ | | ✓ | | ✓ |
| Nationwide identifier does not exist, but social security number can be used under certain constraints | | ✓ | | | ✓ | | ✓ | |
| Not applicable | ✓ | | | | | | | |

## Modes of making data available

Health data is generally made available in pseudonymised format, often in a secure processing environment. When data is anonymized, some of those nodes, such as Germany and Finland, authorise their download.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Secure processing environment for pseudonymised data | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Secure processing environment for fully identifiable data in rare cases | | | | ✓ | | | ✓ | |
| Download of pseudonymised in certain cases | | | ✓ | | | | | |
| Download of anonymous data | | ✓ (for aggregated results) | | ✓ | | | ✓ | |
| Pseudonymisation/anonymisation terminology follows GDPR | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| TTP involved | | ✓ | | | | ✓ | | |

Joint controllership is the exception. Mostly the data are made available in a controller to controller relationship.

### Time limits to make data available

In Finland and Norway, there are time limits for providing the data to the researcher (60 working days for combined data). In Denmark the aim is to provide data on an individual level within 30 days in average. After 3 weeks after application has been admitted, the researcher will get feedback.

In practice, it usually takes longer. In France and Belgium, it usually takes six months to access the data after the permit has been delivered. In Norway, depending on the complexity of the request, it can take up to one year.

### Penalties for health data misuse

Almost all Nodes involve their data protection authorities somehow in the application process, some directly requiring their approval before granting access (France) others as supervisory authorities.

The most common means to react to the misuse of data is exclusion from further applications followed by contractual/administrative penalties.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Exclusion from further applications | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Depends on actual data controller | | | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Criminal penalties | | | | ✓ | ✓ | | ✓ | |

## Involvement of citizens

Citizens are informed about the use of their health data either collectively (website) or individually.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Collective information (website) | | ✓ | | ✓ | ✓ | | ✓ | ✓ |
| Individual information when data are included in databases | | | | | ✓ | | | |
| No information at all | ✓ | | ✓ | | | ✓ | | |

The involvement of citizens in the use of their data does not follow general principles, but ranges from informed consent, opt-out to no involvement at all.

| | BBMRI-ERIC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| General consent requirement | | | | | | | ✓ | |
| General opt out with minor exceptions | | | | ✓ | ✓ | | ✓ | |
| No citizen involvement at all | | ✓ | | | | ✓ | | ✓ |

Furthermore, when it comes to information and transparency, France, Belgium and Norway maintain a public register of projects and a public tool for the researchers to display their results (in development in Germany). Although there is no official register for these in Finland, all the permits issued by Findata are listed in Findata website. Finally, many measures to exercise GDPR rights are still being implemented or even debated.