European Rare Diseases Research Alliance Data Hub

First published: 01/02/2024

Last updated: 17/10/2024

Data source | Human | Hospital outpatient visit records | Hospital inpatient records | Other

Administrative details

Administrative details

Data source ID

1111220

Data source acronym

ERDERA-DH

Data holder

INSERM

Data source type

Hospital outpatient visit records

Hospital inpatient records

Other

Data source type, other

The Data Hub is sourced from partner resource's metadata (now: EJP RD, future: ERDERA and via a governance model), encompassing disease registries that collect common data elements, cell and sample resources, and public data and knowledge resources.

Main financial support

European public funding

Other

Care setting

Hospital outpatient care

Hospital inpatient care

Data source qualification

If the data source has successfully undergone a formal qualification process (e.g., from the EMA, ISO or other certifications), this should be described.

No

Data source website

https://ejprarediseases.org/

Contact details

Daria Julkowska daria.julkowska@inserm.fr

Main

daria.julkowska@inserm.fr

Data source regions and languages

Data source countries

Austria

Belgium

Bulgaria

Croatia

Cyprus

Czechia

Denmark

Estonia

Finland

France

Germany

Greece

Hungary

Iceland

Ireland

Italy

Latvia

Liechtenstein

Lithuania

Luxembourg

Malta

Netherlands

Norway

Poland

Portugal

Romania

Slovakia

Slovenia

Spain

Sweden

United Kingdom

Data source establishment

Data source established

15/06/2021

Data source time span

First collection: 15/06/2021

The date when data started to be collected or extracted.

Publications

Data source publications

FAIR Data Point: A FAIR-Oriented Approach for Metadata Publication

Towards FAIRification of sensitive and fragmented rare disease patient data: challenges and solutions in European reference network registries

Semantic modelling of common data elements for rare disease registries, and a prototype workflow for their deployment over registry data

Enabling FAIR Discovery of Rare Disease Digital Resources

Leveraging Biolink as a FAIR “Rosetta Stone” Between Clinical Semantic Models Provides Emergent Interoperability

Data elements collected

The data source contains the following information

Disease information

Does the data source collect information with a focus on a specific disease? This might be a patient registry or other similar initiatives.

No

Rare diseases

Are rare diseases captured? In the European Union a rare disease is one that affects no more than 5 people in 10,000.

Yes

Pregnancy and/or neonates

Does the data source collect information on pregnant women and/or neonatal subpopulation (under 28 days of age)?

Yes

Hospital admission and/or discharge

No

ICU admission

Is information on intensive care unit admission available?

No

Cause of death

Captured

Cause of death vocabulary

ICD-10

Prescriptions of medicines

Captured

Prescriptions vocabulary

ATC

other

Prescriptions vocabulary, other

See https://github.com/CARE-SM/CARE-Semantic-Model/wiki/CARE-SM-Medication and references therein

Dispensing of medicines

Captured

Dispensing vocabulary

ATC

other

Advanced therapy medicinal products (ATMP)

Is information on advanced therapy medicinal products included? A medicinal product for human use that is either a gene therapy medicinal product, a somatic cell therapy product or a tissue engineered products as defined in Regulation (EC) No 1394/2007 [Reg (EC) No 1394/2007 Art 1(1)].

Yes

Contraception

Is information on the use of any type of contraception (oral, injectable, devices etc.) available?

No

Indication for use

Does the data source capture information on the therapeutic indication for the use of medicinal products?

Captured

Indication vocabulary

Orphacode

Orphanet Rare Disease Ontology (ORDO)

Other

Indication vocabulary, other

NCIT

Medical devices

Is information on medicinal devices (e.g., pens, syringes, inhalers) available?

No

Administration of vaccines

No

Procedures

Does the data source capture information on procedures (e.g., diagnostic tests, therapeutic, surgical interventions)?

Captured

Procedures vocabulary

OPCS

Other

Procedures vocabulary, other

NCIT

Healthcare provider

Is information on the person providing healthcare (e.g., physician, pharmacist, specialist) available? The healthcare provider refers to individual health professionals or a health facility organisation licensed to provide health care diagnosis and treatment services including medication, surgery and

medical devices.

No

Clinical measurements

Is information on clinical measurements (e.g., BMI, blood pressure, height) available?

Yes

Genetic data

Are data related to genotyping, genome sequencing available?

Captured

Genetic data vocabulary

HGNC

Other

Genetic data vocabulary, other

See https://github.com/CARE-SM/CARE-Semantic-Model/wiki/CARE-SM-Genotype and references therein

Biomarker data

Does the data source capture biomarker information? The term “biomarker” refers to a broad subcategory of medical signs (objective indications of medical state observed from outside the patient), which can be measured accurately and reproducibly. For example, haematological assays, infectious disease markers or metabolomic biomarkers.

Not Captured

Patient-reported outcomes

Is information on patient-reported outcomes (e.g., quality of life) available?

Yes

Patient-generated data

Is patient-generated information (e.g., from wearable devices) available?

Yes

Units of healthcare utilisation

Are units of healthcare utilisation (e.g., number of visits to GP per year, number of hospital days) available or can they be derived? Units of healthcare utilisation refer to the quantification of the use of services for the purpose of preventing or curing health problems.

No

Unique identifier for persons

Are patients uniquely identified in the data source?

No

Diagnostic codes

Captured

Diagnosis / medical event vocabulary

Orphacode

Orphanet Rare Disease Ontology (ORDO)

Other

Diagnosis / medical event vocabulary, other

NCIT and others. See https://github.com/CARE-SM/CARE-Semantic-Model/wiki/CARE-SM-Diagnosis and references therein

Medicinal product information

Captured

Medicinal product information collected

Active ingredient(s)

Medicinal product vocabulary

ATC

Other

If 'other,' what vocabulary is used?

See https://github.com/CARE-SM/CARE-Semantic-Model and references therein

Quality of life measurements

Captured

Quality of life measurements vocabulary

SF-36

EQ5D

other

Quality of life measurements, other

Other also being/will be collected (e.g., PedSQL)

Lifestyle factors

Captured

Lifestyle factors

Other

Sociodemographic information

Captured

Sociodemographic information collected

Age

Gender

Country of origin

Quantitative descriptors

Population age groups

Paediatric Population (< 18 years)

Preterm newborn infants (0 – 27 days)

Term newborn infants (0 – 27 days)

Infants and toddlers (28 days – 23 months)

Children (2 to < 12 years)

Adolescents (12 to < 18 years)

Adults (18 to < 46 years)

Adults (46 to < 65 years)

Elderly (≥ 65 years)

Adults (65 to < 75 years)

Adults (75 to < 85 years)

Adults (85 years and over)

Estimated percentage of the population covered by the data source in the catchment area

Increasing coverage, working towards inclusion of data from all 24 European Reference Networks on Rare Diseases.

Population

Population size

55000

Median observation time

Median time (years) between first and last available records for unique individuals captured in the data source

2.50

Data flows and management

Biospecimen access

Are biospecimens available in the data source (e.g., tissue samples)?

Yes

Biospecimen access conditions

Availability and accessibility depends on the source registry

Access to subject details

Can individual patients/practitioners/practices included in the data source be contacted?

No

Description of data collection

Reference Centers for specific rare diseases provide data to the registries of the ERN (see #31). Data of these registries is standardized and made FAIR, and made available for access through the virtual platform.

Event triggering registration

Event triggering registration of a person in the data source

Disease diagnosis

Event triggering de-registration of a person in the data source

Loss to follow up

Other

Event triggering de-registration of a person in the data source, other

Consent retraction

Event triggering creation of a record in the data source

Currently via a collaborative onboarding process. ERNs propose the enrolment of patients in the registry whenever diagnostic is established

Data source linkage

Linkage

Is the data source described created by the linkage of other data sources (prelinked data source) and/or can the data source be linked to other data source on an ad-hoc basis?

Yes

Linkage description, pre-linked

See https://specs.fairdatapoint.org/

https://github.com/ejp-rd-vp/resource-metadata-schema

1. Metadata in RDF conform to the W3C-recommended DCAT model (currently version 2.0)
2. A discovery API conform to the GA4GH Beacon-2 framework schema and API
3. Data element metadata conform to an ontology-based common data element model (https://github.com/CARE-SM/CARE-Semantic-Model)

Linked data sources

Pre linked

Is the data source described created by the linkage of other data sources?

Yes

Data source, other

Each linked data source provides its name through FAIR metadata: resources are 'linked' via a URI, referencing a locally deployed metadata provisioning service that gives access through http to metadata defined in terms of the Data

Catalog Vocabulary and extensions thereof, conform the FAIR Data Point specifications (https://specs.fairdatapoint.org/)

Data management specifications that apply for the data source

Informed consent for use of data for research

Waiver

Possibility of data validation

Can validity of the data in the data source be verified (e.g., access to original medical charts)?

Yes

Data source preservation

Are records preserved in the data source indefinitely?

No

Approval for publication

Is an approval needed for publishing the results of a study using the data source?

No

Common Data Model (CDM) mapping

CDM mapping

Has the data source been converted (ETL-ed) to a common data model?

Yes

CDM Mappings

CDM name (other)

Clinical And Registry Entries Semantic Model (CARE-SM)

Data source ETL specifications (link)

https://github.com/ejp-rd-vp/CARE-Semantic-Model

CDM name

OMOP

CDM website

https://www.ohdsi.org/Data-standardization/

Data source ETL CDM version

4.5.1

Data source ETL status

In progress