Establish an EU catalogue of sources of real-world data, characterised by a common set of metadata and data quality measurements

First published: 05/12/2022

Last updated: 12/06/2024

Study **Finalised**

Administrative details

EU PAS number

EUPAS49303

Study ID

49304

DARWIN EU® study

No

Study countries

☐ Belgium
☐ Czechia
☐ Denmark

☐ Finland

☐ France

☐ Germany

☐ Italy

☐ Netherlands

☐ Norway

☐ Poland

☐ Slovakia

☐ Spain

☐ Sweden

☐ United Kingdom

Study description

Europe is generating unprecedented amounts of person-level information contained in Electronic Medical Record (EMR) systems. This includes structured data in the form of diagnoses, medication, and laboratory test results. These EMR datasets are often siloed by country, language, region, hospital and even department, captured in a disease specific context, however, these data contain invaluable insights for regulatory purposes. Metadata is traditionally defined as “data about data”. It’s a set of data that describes and gives information on other data providing context about their purpose, location, key-variables, generation, format, and ownership of a dataset. Metadata help both researchers and study sponsors to identify datasets for conducting specific studies. Currently, publicly available metadata is limited, not coherent between different sources, and not defined with a regulatory purpose in mind. Therefore, the Agency is currently developing an EU catalogue of real-world data sources which will replace the existing ENCePP catalogue in late 2023. To accommodate this process, the Agency awarded IQVIA to propose a methodology to grow and maintain the catalogue in a sustainable manner, as well as to catalogue an initial set of 24 data sources in 2022. The overall objective is to establish an EU catalogue of sources of real-world data, characterised by a common set of metadata and data quality measurements.

Study status

Finalised

Research institutions and networks

Institutions

IQVIA NL, Real-World-Evidence

☐ Netherlands

First published: 25/11/2022

Last updated: 21/03/2025

Institution **Other** **ENCePP partner**

Multiple centres: 24 centres are involved in the study

Contact details

Study institution contact

Michèle Arnoe michele.arnoe@iqvia.com

Study contact

michele.arnoe@iqvia.com

Primary lead investigator

Michèle Arnoe

Primary lead investigator

Study timelines

Date when funding contract was signed

Planned: 13/12/2021

Actual: 13/12/2021

Study start date

Planned: 01/07/2022

Actual: 01/07/2022

Date of final study report

Planned: 23/01/2023

Actual: 21/03/2023

Sources of funding

- EMA

Regulatory

Was the study required by a regulatory body?

Yes

Is the study required by a Risk Management Plan (RMP)?

Not applicable

Methodological aspects

Study type

Study type list

Study topic:

Other

Study topic, other:

Disease/Epidemiology study

Study type:

Not applicable

Scope of the study:

Other

If ‘other’, further details on the scope of the study

Propose a methodology to grow the EU catalogue of real-world data sources in a sustainable manner and catalogue an initial set of 24 data sources.

Data collection methods:

Combined primary data collection and secondary use of data

Main study objective:

To establish an EU catalogue of sources of real-world data, characterised by a common set of metadata and data quality measurements.

Population studied

Short description of the study population

N/A

Age groups

- 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)
- Adults (65 to < 75 years)
- Adults (75 to < 85 years)
- Adults (85 years and over)

Estimated number of subjects

0

Study design details

Data analysis plan

Not applicable

Data management

ENCePP Seal

The use of the ENCePP Seal has been discontinued since February 2025. The ENCePP Seal fields are retained in the display mode for transparency but are no longer maintained.

Data sources

Data source(s)

BIFAP - Base de Datos para la Investigación Farmacoepidemiológica en el Ámbito Público (Pharmacoepidemiological Research Database for Public Health Systems)

Pedianet

PedNet Haemophilia registry

Hospital Episode Statistics

Sweden National Prescribed Drugs Register / Läkemedelsregistret

Danish registries (access/analysis)

European network of population-based registries for the epidemiological surveillance of congenital anomalies

---

Data sources (types)

Administrative healthcare records (e.g., claims)

Disease registry

Drug dispensing/prescription data

Electronic healthcare records (EHR)

Other

Data sources (types), other

Prospective patient-based data collection

Use of a Common Data Model (CDM)

CDM mapping

No

Data quality specifications

Check conformance

Unknown

Check completeness

Unknown

Check stability

Unknown

Check logical consistency

Unknown

Data characterisation

Data characterisation conducted

No