Establish an EU catalogue with metadata of data sources and studies

First published: 31/03/2023

Last updated: 23/09/2024

Study Finalised

Administrative details

EU PAS number

EUPAS104093

Study ID

104094

DARWIN EU® study

No

Study countries

- Belgium
- France
- Germany
- Netherlands
- Portugal

Study description

The identification of appropriate data sources to answer specific research questions and to assess the suitability of these data sources for a study might be challenging. To support the identification of suitable data sources, the HMA-EMA joint Big Data Task Force recommended “to promote data discoverability through the identification of metadata” as part of its Recommendation III, which is reflected in the Big Data Steering Group work plan. Metadata are descriptive data that characterise other data to create a better understanding of their meaning and to achieve greater reliability and quality when using data for a specific purpose. The Agency is developing an electronic catalogue that will provide metadata for data sources and studies targeted for release in late 2023. The Agency started collecting metadata of data sources in 2022 and an initial set of 24 data sources was captured (EU PAS EUPAS49303). To further accommodate this process, the Agency started a process with IQVIA to collect a defined set of metadata for 50 data sources in 2023 and a process to verify and refine the defined set of metadata information for 1,270 finalised studies migrated from EU PAS Register® to the new studies catalogue. Main objective of this study: • The overall objective is to establish an EU catalogue of data sources and studies, characterised by a common set of metadata and data quality measurements. Specific objectives are: • To verify and refine the defined set of metadata for finalised studies migrated from the current EU PAS Register® to the new catalogue of studies. • To collaborate with data holders in order to collect the defined set of metadata for data sources migrated from the ENCePP resource database to the new data sources catalogue and also other data sources. This includes quality checking of the information received from data holders.

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

Contact details

Study institution contact

Hanne van Ballegooijen hanne.vanballegooijen@iqvia.com

Study contact

hanne.vanballegooijen@iqvia.com

Primary lead investigator

Michèle Arnoe

Primary lead investigator

Study timelines

Date when funding contract was signed

Planned: 05/12/2022

Actual: 05/12/2022

Study start date
Planned: 01/03/2023

Actual: 22/03/2023

Date of final study report
Planned: 15/12/2023

Actual: 01/07/2024

Sources of funding

- EU institutional research programme

More details on 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 type:

Not applicable

Scope of the study:

Drug utilisation

Other

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

EMA Big Data initiative

Main study objective:

EMA requires additional work related to the enrichment of data sources metadata following the migration of the current ENCePP resource database to the new catalogue. EMA has mapped the data fields between the ENCePP Resource Database and the new studies catalogue. IQVIA will provide support to the data sources for the completion of the spreadsheet, will quality check the file and provide feedback.

Population studied

Short description of the study population

This project is not a study but rather part of the EMA Big Data initiative to establish an EU catalogue of real-world data sources. Any responses regarding “data” in this questionnaire refer to metadata (not individual-level data) for the purposes of this project.

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

1000000

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), other

Multiple sources were included in the HMA-EMA Catalogues. Examples of data sources: Electronic Health records, Claims, Drug prescription, Disease registry, Administrative data.

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, Exposure registry, Case-control surveillance database

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