

# Strengthening Use of Real-World Data in Medicines Development: Metadata for Data Discoverability and Study Replicability (MINERVA)

**First published:** 03/02/2021

**Last updated:** 16/12/2024

Study

Finalised

## Administrative details

### EU PAS number

EUPAS39322

---

### Study ID

49345

---

### DARWIN EU® study

No

---

### Study countries

 Austria

 Belgium

 Czechia

-  Denmark
  -  Estonia
  -  France
  -  Germany
  -  Italy
  -  Netherlands
  -  Slovenia
  -  Spain
  -  Sweden
  -  Switzerland
  -  United Kingdom
- 

## **Study description**

Identification of appropriate data sources is becoming an increasing need for regulatory decision making. Metadata are descriptive data that characterise other data to create a clearer understanding of their meaning and achieve greater reliability and quality of information. Access to a standard and electronic set of complete and accurate metadata information can contribute to identifying the datasources that are most suitable for a specific study, describe the datasources planned to be used in a study protocol or research proposal, and contribute to assessing the evidentiary value of the results of a multidatabase observational study submitted in a study report or publication. The objective of this study is to define and collect a set of metadata for a selected number of real-world (RW) datasources. The study consists in the preparatory step for the development and population of a publicly available catalogue of sources of RW data characterised with a common set of metadata. The catalogue will serve as a proof of concept for a metadata inventory with searchability, dynamic dashboards, visualisation capabilities, and integration of different levels of access (public vs. confidential information). Specific objectives of this study are:

1. Define a list of criteria to identify relevant RW datasources from which the datasources to be included in this study will be selected.
  2. Identify a list of a minimum of 10 datasources for this study.
  3. Define a set of metadata that should be collected from RW datasources.
  4. Conduct an in-depth stakeholders' consultation on the metadata identified.
  5. Define a process to collect the set of metadata for the datasources in the study.
  6. Collect the defined set of metadata for datasources in the study.
  7. Develop a tool enabling access to the metadata collected.
  8. Draft a good practice guide with the description of the metadata defined and recommendations on its use for the purpose of identifying RW datasources for specific studies.
- 

## Study status

Finalised

# Research institutions and networks

## Institutions

### RTI Health Solutions (RTI-HS)

-  France
-  Spain
-  Sweden
-  United Kingdom
-  United Kingdom (Northern Ireland)
-  United States

**First published:** 21/04/2010

**Last updated:** 13/03/2025

**Institution**

**Not-for-profit**

**ENCePP partner**

## Electronic Health Records (EHR) Research Group, London School of Hygiene & Tropical Medicine (LSHTM)

 United Kingdom

**First published:** 19/04/2010

**Last updated:** 30/10/2024

**Institution**

**Educational Institution**

**ENCePP partner**

## Division of Pharmacoepidemiology & Clinical Pharmacology (PECP), Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University

 Netherlands

**First published:** 01/03/2010

**Last updated:** 23/05/2024

**Institution**

**Educational Institution**

**ENCePP partner**

## Health Services Research and Pharmacoepidemiology Unit (HSRP Unit) FISABIO

 Spain

**First published:** 30/11/2023

**Last updated:** 30/11/2023

**Institution**

Other

ENCePP partner

## University Medical Center Utrecht (UMCU)

 Netherlands

**First published:** 24/11/2021

**Last updated:** 22/02/2024

**Institution**

Educational Institution

Hospital/Clinic/Other health care facility

ENCePP partner

## Centre for Pharmacoepidemiology, Karolinska Institutet (CPE-KI)

 Sweden

**First published:** 24/03/2010

**Last updated:** 23/04/2024

**Institution**

Educational Institution

Laboratory/Research/Testing facility

Not-for-profit

ENCePP partner

## Bordeaux PharmacoEpi, University of Bordeaux

 France

**First published:** 07/02/2023

**Last updated:** 08/12/2025

**Institution**

Educational Institution

Hospital/Clinic/Other health care facility

Not-for-profit

ENCePP partner

## University of Tartu

 Estonia

**First published:** 01/02/2024

**Last updated:** 01/02/2024

**Institution**

Educational Institution

## Agenzia regionale di sanità della Toscana (ARS Toscana)

 Italy

**First published:** 01/02/2024

**Last updated:** 23/03/2026

**Institution**

**EU Institution/Body/Agency**

**ENCePP partner**

## Agencia Española de Medicamentos y Productos Sanitarios (Spanish Agency for Medicines and Medical Devices, AEMPS)

 Spain

**First published:** 01/02/2024

**Last updated:** 04/09/2024

**Institution**

**EU Institution/Body/Agency**

**Not-for-profit**

**Regulatory Authority**

**ENCePP partner**

## European Society for Blood and Marrow Transplantation (EBMT)

 European Union

**First published:** 20/02/2024

**Last updated:** 20/02/2024

**Institution**

University Medical Center Groningen (The Netherlands), The Spanish Working Group on

Chron's Disease and Ulcerative Colitis (GETECCU)  
(Spain), University of Ljubljana Faculty of  
Pharmacy (Slovenia), Institute of Genomics  
(Estonia)

## Networks

### The SIGMA Consortium (SIGMA)

-  Denmark
-  European Union
-  France
-  Germany
-  Italy
-  Netherlands
-  Norway
-  Spain
-  Sweden
-  United Kingdom

**First published:** 10/02/2013

**Last updated:** 19/01/2026

Network

ENCePP partner

# EU Pharmacoepidemiology and Pharmacovigilance (PE&PV) Research Network

 Netherlands

**First published:** 01/02/2024

**Last updated:** 24/09/2025

Network

## Contact details

### Study institution contact

Susana Perez-Gutthann [sperez@rti.org](mailto:sperez@rti.org)

Study contact

[sperez@rti.org](mailto:sperez@rti.org)

### Primary lead investigator

Susana Perez-Gutthann 0000-0001-5798-3691

Primary lead investigator

### ORCID number:

0000-0001-5798-3691

## Study timelines

### Date when funding contract was signed

Planned: 04/01/2021

Actual: 23/11/2020

---

**Study start date**

Planned: 31/05/2021

Actual: 01/03/2021

---

**Date of final study report**

Planned: 15/12/2021

Actual: 10/01/2022

## 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:**

Infrastructure and guidance development

**Study type:**

Not applicable

---

**If 'other', further details on the scope of the study**

Definition and description of metadata, visualisation catalogue proof of concept tool and development of a good practice guide

**Main study objective:**

To define the set of metadata that should be collected from RW data sources and to collect the metadata for at least 10 data sources for the future development & population of a public catalogue with searchability, dashboards, visualisation capabilities, and integration of different levels of access

## Population studied

**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)

## Study design details

### Data analysis plan

Not applicable

## Documents

### Study results

[MINERVA\\_GoodPracticeGuide\\_10Jan2022.pdf](#) (1.8 MB)

---

### Study report

[MINERVA\\_D4\\_WorkshopReport\\_Final.pdf](#) (764.56 KB)

[MINERVA\\_DataSourceSelection\\_Final\\_26March2021.pdf](#) (1.17 MB)

### Study, other information

[MINERVA Full-metadata-list 10Jan2022.pdf](#) (968.07 KB)

[MINERVA\\_D5\\_Report\\_v1.1\\_28June2021.pdf](#) (2.46 MB)

[MINERVA\\_D9\\_Metadata-catalogue-10Jan2022.pdf](#) (587.5 KB)

[MINERVA\\_DataSourceSelection\\_Final\\_26March2021.pdf](#) (1.17 MB)

### Study publications

[Pajouheshnia, R., Gini, R., Gutierrez, L., Swertz, M., Hyde, E., Sturkenboom, M...](#)

[Gini, R., Pajouheshnia, R., Gutierrez, L., Swertz, M., Hyde, E., Sturkenboom, M...](#)

[Pajouheshnia R, Gini R, Hyde E, Swertz MA, Sturkenboom M, Margulis AV,](#)

[Franzoni...](#)

## 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)**

Clinical Practice Research Datalink

Danish registries (access/analysis)

PHARMO Data Network

German Pharmacoepidemiological Research Database

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

ARS Toscana

EpiChron Cohort

The Valencia Health System Integrated Database

Système National des Données de Santé (French national health system main database)

**Data source(s), other**

EBMT; Swedish National Registers, Sweden; Slovenian health data, Slovenia.

---

**Data sources (types)**

Administrative healthcare records (e.g., claims)

Disease registry

Drug dispensing/prescription data

Electronic healthcare records (EHR)

Laboratory tests and analyses

Omics

Population registry

Pregnancy registry

## Use of a Common Data Model (CDM)

**CDM mapping**

No

## Data quality specifications

**Check conformance**

Yes

---

**Check completeness**

Yes

---

**Check stability**

Yes

---

**Check logical consistency**

Yes

## Data characterisation

**Data characterisation conducted**

Not applicable