Describing diversity of real world data sources in pharmacoepidemiologic studies: The DIVERSE scoping review and tool

First published: 04/03/2021

Last updated: 26/08/2026

Study Ongoing

Administrative details

EU PAS number

EUPAS39757

Study ID

49419

DARWIN EU® study

No

Study countries

Italy

Netherlands

United Kingdom (Northern Ireland)

Study description

This study addresses reproducibility challenges associated with diversity across real-world data sources repurposed for secondary use in pharmacoepidemiologic studie. The DIVERSE project is conducted by the Database Special Interest Group of the International Society for Pharmacoepidemiology (ISPE). It encompassises two objectives: Objective 1 was conducting a scoping review and Objective 2 was developing a tool to represent data diversity.

Study status

Ongoing

Research institutions and networks

Institutions

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**

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

Netherlands

First published: 01/03/2010

Last updated: 27/05/2026

Institution | Educational Institution | ENCePP partner

Contact details

Study institution contact

Rosa Gini rosa.gini@ars.toscana.it

Study contact

rosa.gini@ars.toscana.it

Primary lead investigator

Rosa Gini

Primary lead investigator

Study timelines

Date when funding contract was signed

Planned: 24/12/2020

Actual: 24/12/2020

Study start date

Planned: 01/01/2021

Actual: 01/01/2021

Date of final study report

Planned: 30/09/2027

Sources of funding

- Other

More details on funding

International Society for Pharmacoepidemiology (ISPE)

Study protocol

DIVERSE_protocol_v1.0.pdf (280.14 KB)

Regulatory

Was the study required by a regulatory body?

No

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:

Methodology to generate Real World Evidence

Study type:

Not applicable

Scope of the study:

Method development or testing

Scoping review (including literature review)

Study design:

scoping review

Main study objective:

To list and summarize existing tools and recommendations for the collection and reporting of diversity in data sources used in RWE. To develop a tool to represent data diversity.

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)

Estimated number of subjects

0

Study design details

Data analysis plan

In the scoping review, articles were identified through searches in PubMed and Embase and by expert knowledge of relevant literature. The sensitivity of the search was validated against a set of relevant papers identified by experts in the working group. Articles were screened on title and abstract in duplicate using a standard screening tool and will be included if they contain recommendations or guidelines for the collection and reporting of (diversityof) data sources, report tools to describe data sources or provide descriptions of multiple data sources within a network. Information was collected from the selected articles using a data extraction tool, applied in duplicated by two independent researchers. Extracted information was analysed in accordance with recommendations in the JBI Manual for Evidence Synthesis, and will follow the Arksey and O'Malley framework for collating and summarizing results in a narrative review. Counts of different types of articles was provided.

Summary results

Scoping review: Across 91 selected documents, 9 dimensions used to describe Real World Data Sources (organization accessing the data source, data

originator, prompt, inclusion of population, content, data dictionary, time span, healthcare system and culture, and data quality). Thirty-six themes were identified within the dimensions

Tool: available at this link https://docs.google.com/spreadsheets/d/1yRCmjb71smATl5bk1yJTNcy-o4HyKZfD/edit?gid=428534828#gid=428534828 and as a supplement in this Catalogues entry, in section Study Documents

Documents

Study results

The DIVERSE Tool.pdf (41.41 KB)

First peer-reviewed publication from the study (scoping review)

Study, other information

poster_DIVERSE_rev.pdf (370.01 KB)

DIVERSE study supplementary materials - supplementary materials to the manuscri…

The DIVERSE tool (to be downloaded and used)

Study publications

Published scoping review

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 sources (types)

Other

Published literature

Data sources (types), other

Electronic medical literature databases: PubMed and Embase

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