Data source heterogeneity in multi-database pharmacoepidemiologic studies: a scoping review (DIVERSE)

First published: 04/03/2021

Last updated: 02/07/2024

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

Multi-database studies (MDS) are increasingly performed in pharmacoepidemiologic research. A MDS is as a study using at least two healthcare databases, which are not linked with each other at an individual person level, either because they cover and capture information on different individuals, or because, even if populations overlap, local regulations forbid record linkage. In a MDS, analyses are carried out in parallel across each data source applying a common study protocol. Regulatory authorities often require data from multiple data sources to be used in a single study, to enhance the generalizability of results or to obtain sufficient sample size when the exposure and/or outcome is rare. MDS pose a number of challenges, including how to manage heterogeneity between the different included data sources. Despite calls for the implementation of strategies to improve replicability, increase transparency and reduce bias in MDS, and despite general recommendations to assess the comparability of data sources in MDS, to our knowledge, there is currently no guidance for how database heterogeneity should be evaluated or even identified and recorded. This scoping review is intended to inform the development of guidelines for the identification, collection and reporting of heterogeneity in MDS, and to identify areas for further research. This activity is the Objective 1 of the DIVERSE project, of the Database Special Interest Group of the International Society for Pharmacoepidemiology (ISPE).

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: 31/12/2023

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 type:

Not applicable

Main study objective:

To list and summarize existing tools and recommendations for the collection and reporting of heterogeneity in data sources used in MDS, in particular listing and classifying existing descriptors of such heterogeneity. A secondary objective is to describe how heterogeneity is leveraged to improve the quality of the evidence generated in a MDS and to assist its interpretation.

Population studied

Age groups

- Adolescents (12 to < 18 years)

- Children (2 to < 12 years)
- Infants and toddlers (28 days – 23 months)
- Preterm newborn infants (0 – 27 days)
- Term newborn infants (0 – 27 days)
- 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

Articles will be identified through searches in PubMed and Embase and by expert knowledge of relevant literature. The sensitivity of the search will be validated against a set of relevant papers identified by experts in the working group. Articles will be 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 (heterogeneity of) data sources, report tools to describe data sources or provide descriptions of multiple data sources within a network. Information will be collected from the selected articles using a data extraction tool, applied in duplicated by two independent researchers. Extracted information will be 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 will be provided.

Documents

Study, other information

DIVERSE_selection_tool_fulltext_final_pdf.pdf (123.13 KB)

DIVERSE_selection_tool_TIAB_final_pdf.pdf (93.35 KB)

poster_DIVERSE_rev.pdf (370.01 KB)

Report on DIVERSE Task 1a2_v1.1.pdf (1.02 MB)

Report on DIVERSE Task 1a4_v1.11.pdf (1.54 MB)

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