

# The Valencia Health System Integrated Database

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

**Last updated:** 08/06/2026

Data source

Human

Hospital outpatient visit records

Other

Primary care medical records

Registration with healthcare system

## Administrative details

### Administrative details

#### Data source ID

1111174

---

#### Data source acronym

VID

---

#### Data holder

[The Foundation for the Promotion of Health and Biomedical Research of Valencia Region \(FISABIO\)](#)

---

#### Data source type

Hospital outpatient visit records

Other

Primary care medical records  
Registration with healthcare system

---

### **Data source type, other**

VID also contains hospital discharge records, emergency care discharge records, birth registry, congenital anomaly registry, perinatal mortality registry, pharmacy prescription and dispensing records, vaccine records, and microbiology records .

---

### **Main financial support**

European public funding  
Funding by own institution

---

### **Care setting**

Hospital inpatient care  
Hospital outpatient care  
Primary care – GP, community pharmacist level  
Primary care – specialist level (e.g. paediatricians)  
Secondary care – specialist level (ambulatory)

---

### **Data source qualification**

If the data source has successfully undergone a formal qualification process (e.g., from the EMA, ISO or other certifications), this should be described.

No

## Contact details

Gabriel Sanf elix-Gimeno [gabriel.sanfelix@fisabio.es](mailto:gabriel.sanfelix@fisabio.es)

Main

## Data source regions and languages

### Data source countries

Spain

---

### Data source languages

Catalan

Spanish

---

### Data source regions

Valenciana, Comunidad

## Data source establishment

### Data source established

01/01/2008

---

### Data source time span

**First collection:** 01/01/2008

The date when data started to be collected or extracted.

## Publications

### Data source publications

[Data Resource Profile: The Valencia Health System Integrated Database \(VID\)](#)

## Studies

# List of studies that have been conducted using the data source

CONSIGN study: COVID-19 infection and medicines in pregnancy - a multinational registry based study

A post-authorisation/post-marketing observational study to evaluate the association between exposure to AZD1222 and safety concerns using existing secondary health data sources (COVID-19)

An Observational Post-Authorization Safety Study to Assess the Safety of Ad26.COV2.S Using European Healthcare Data through VAC4EU (COVID-19)

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

mRNA-1273-P910: Clinical course, outcomes and risk factors of myocarditis and pericarditis following administration of Moderna vaccines targeting SARS-CoV-2.

SAFETY-VAC: Network of Data Sources for Vaccine Safety Evaluation

ADEPT: The utilisation of antiseizure medications in pregnant women, other women of childbearing potential, and men: a multi-database study from 7 European countries

SAFETY-VAC: Background incidence estimation of flares of pre-existing chronic diseases using pan-European electronic healthcare data sources.

DARWIN EU® - CGRP antagonists - Treatment patterns and users characteristics

ADEPT: feasibility of estimating the risk of adverse pregnancy, neonatal and child outcomes following either in utero ASM exposure through the mother, or peri-conceptual ASM exposure through the father

SAFETY-VAC: Phenotype proposal and rates of immunocompromised populations in real-world data sources.

DARWIN EU® - Incidence of suicidality in patients with specific chronic skin conditions

VAC4EU Postauthorisation Safety Study of BIMERVAX® Vaccine in Europe

VAC4EU Postauthorisation Effectiveness Study of BIMERVAX® Vaccine in Europe

A Post-Authorisation Safety Study (PASS) of ABRYSSVO (Respiratory Syncytial Virus Stabilised Prefusion Subunit Vaccine) in Pregnant Women and their Offspring in a Real World Setting in Europe and UK (C3671026)

A post-authorisation safety study of ABRYSSVO in immunocompromised, or renally or hepatically impaired adults aged 60 years and older in a real world setting in Europe and UK (C3671038)

DARWIN EU® - Characterising the use of JAK inhibitors in Europe: a Drug Utilisation Study

TARGET EU: Comparative effectiveness and safety studies using the target trial emulation and estimand frameworks

Use of antipsychotics in children and adolescents in the Spanish National Health System: network project with real-world data from five Autonomous Communities to inform evidence-based health policies (Children and Adolescents Antipsychotic Research for Evidence-based Strategies - CARES Project)

DARWIN EU® - Capturing suicidality and depression related variables in databases

DARWIN EU® - Characteristics of individuals with acute graft vs. host disease with intestinal involvement

Estimation of background incidence rates of coagulation disorders and the association with COVID-19 vaccines in pregnant population: a multi-database study from 3 European countries

DARWIN EU® - Population demographics and disease frequency across the DARWIN EU® network

DARWIN EU® - Characterising the incidence and presentation of myocarditis and/or pericarditis in Europe

TARGET-EU: Effectiveness of BNT162b2 mRNA COVID-19 vaccine in healthy individuals or with stable pre-existing medical conditions against SARS-CoV-2 infection

DARWIN EU® - Characterising the use of JAK inhibitors in Europe: a Drug Utilisation Study Update

TARGET-EU: Risk of adverse birth and neurodevelopmental outcomes in children born alive to fathers exposed to valproate versus levetiracetam for generalised epilepsy

## Data elements collected

The data source contains the following information

### **Disease information**

Does the data source collect information with a focus on a specific disease? This might be a patient registry or other similar initiatives.

Yes

---

### **Disease details (other)**

the REDMIVA is the table where microbiological surveillance network information is collected, including the information about COVID-19 test results.

---

### **Rare diseases**

Are rare diseases captured? In the European Union a rare disease is one that affects no more than 5 people in 10,000.

Yes

---

### **Pregnancy and/or neonates**

Does the data source collect information on pregnant women and/or neonatal subpopulation (under 28 days of age)?

Yes

---

### **Hospital admission and/or discharge**

Yes

---

### **ICU admission**

Is information on intensive care unit admission available?

Yes

---

### **Cause of death**

Captured

---

### **Cause of death vocabulary**

ICD-10-CM

---

### **Prescriptions of medicines**

Captured

---

### **Prescriptions vocabulary**

ATC

other

---

### **Prescriptions vocabulary, other**

national vocabulary

---

### **Dispensing of medicines**

Captured

---

### **Dispensing vocabulary**

ATC

other

---

### **Dispensing vocabulary, other**

national vocabulary

---

### **Advanced therapy medicinal products (ATMP)**

Is information on advanced therapy medicinal products included? A medicinal product for human use that is either a gene therapy medicinal product, a somatic cell therapy product or a tissue engineered products as defined in Regulation (EC) No 1394/2007 [Reg (EC) No 1394/2007 Art 1(1)].

No

---

### **Contraception**

Is information on the use of any type of contraception (oral, injectable, devices etc.) available?

Yes

---

### **Indication for use**

Does the data source capture information on the therapeutic indication for the use of medicinal products?

Captured

---

## **Indication vocabulary**

ICD-10-CM

ICD-9-CM

---

## **Medical devices**

Is information on medicinal devices (e.g., pens, syringes, inhalers) available?

Yes

---

## **Administration of vaccines**

Yes

---

## **Procedures**

Does the data source capture information on procedures (e.g., diagnostic tests, therapeutic, surgical interventions)?

Captured

---

## **Procedures vocabulary**

ICD-10

ICD-9

---

## **Healthcare provider**

Is information on the person providing healthcare (e.g., physician, pharmacist, specialist) available?  
The healthcare provider refers to individual health professionals or a health facility organisation licensed to provide health care diagnosis and treatment services including medication, surgery and medical devices.

Yes

---

## **Clinical measurements**

Is information on clinical measurements (e.g., BMI, blood pressure, height) available?

Yes

---

## **Genetic data**

Are data related to genotyping, genome sequencing available?

Not Captured

---

## **Biomarker data**

Does the data source capture biomarker information? The term “biomarker” refers to a broad subcategory of medical signs ( objective indications of medical state observed from outside the patient), which can be measured accurately and reproducibly. For example, haematological assays, infectious disease markers or metabolomic biomarkers.

Not Captured

---

## **Patient-reported outcomes**

Is information on patient-reported outcomes (e.g., quality of life) available?

No

---

## **Patient-generated data**

Is patient-generated information (e.g., from wearable devices) available?

Yes

---

## **Units of healthcare utilisation**

Are units of healthcare utilisation (e.g., number of visits to GP per year, number of hospital days) available or can they be derived? Units of healthcare utilisation refer to the quantification of the use of services for the purpose of preventing or curing health problems.

Yes

---

## **Unique identifier for persons**

Are patients uniquely identified in the data source?

Yes

---

## **Diagnostic codes**

Captured

---

## **Diagnosis / medical event vocabulary**

ICD-10-CM

ICD-9-CM

---

## **Medicinal product information**

Captured

---

## **Medicinal product information collected**

Active ingredient(s)

Dosage regime

Formulation

Package size

Strength

---

## **Medicinal product vocabulary**

ATC

Other

---

## **If 'other,' what vocabulary is used?**

national vocabulary

---

## **Quality of life measurements**

Not Captured

---

## **Lifestyle factors**

Captured

---

## **Lifestyle factors**

Alcohol use

Diet

Frequency of exercise

Tobacco use

---

### **Sociodemographic information**

Captured

---

### **Sociodemographic information collected**

Age

Country of origin

Deprivation index

Gender

Health area

Living in rural area

Pharmaceutical copayment

Socioeconomic status

## Quantitative descriptors

## Population Qualitative Data

### **Population age groups**

Paediatric Population (< 18 years)

Neonate

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)  
Adult and elderly population ( $\geq 18$  years)  
Adults (18 to < 65 years)  
Adults (18 to < 46 years)  
Adults (46 to < 65 years)  
Elderly ( $\geq 65$  years)  
Adults (65 to < 75 years)  
Adults (75 to < 85 years)  
Adults (85 years and over)

---

**Estimated percentage of the population covered by the data source in the catchment area**

97%

---

**Description of the population covered by the data source in the catchment area whose data are not collected (e.g., people who are registered only for private care)**

The data source contains all general population covered by the universal public health care system. About 97% of the population of the region it is covered by the public care

## Family linkage

**Family linkage available in the data source permanently or can be created on an ad hoc basis**

Permanently

---

**Family linkage available between the following persons**

Father-child  
Household

Mother-child

## Population

### Population size

7000000

---

### Active population size

5380000

## Population by age group

Age group	Active population size
Paediatric Population (< 18 years)	830000
Adults (18 to < 46 years)	1900000
Adults (46 to < 65 years)	1500000
Elderly ( $\geq$ 65 years)	1150000

## Median observation time

**Median time (years) between first and last available records for unique individuals captured in the data source**

12.00

---

**Median time (years) between first and last available records for unique active individuals (alive and currently registered) capt**

10.00

Data flows and management

## **Governance details**

Documents or webpages that describe the overall governance of the data source and processes and procedures for data capture and management, data quality check and validation results (governing data access or utilisation for research purposes).

<https://www.san.gva.es/es/web/investigacio/solicitud-datos-sia-gaia>

## **Biospecimen access**

Are biospecimens available in the data source (e.g., tissue samples)?

No

---

## **Access to subject details**

Can individual patients/practitioners/practices included in the data source be contacted?

No

---

## **Description of data collection**

Each table of the data source has their own mechanism of collection and recording of data. Some of the EHR are created by the physicians, some of them by administrators, documentalists, etc.

# Event triggering registration

## **Event triggering registration of a person in the data source**

Birth

Immigration

Other

---

## **Event triggering registration of a person in the data source, other**

Any contact with the health system triggers the registration of the person

---

## **Event triggering de-registration of a person in the data source**

Death

Emigration

Insurance coverage end

---

## **Event triggering creation of a record in the data source**

Most of them are created when a contact with the health system is produced. In other cases, such as pharmacy data, when the prescription is created or when the dispensing is produced. Each table of the data source has their own triggers

## Data source linkage

### **Linkage**

Is the data source described created by the linkage of other data sources (prelinked data source) and/or can the data source be linked to other data source on an ad-hoc basis?

Yes

---

### **Linkage description, pre-linked**

The linkage among all the data sources is produced by an ID pseudonymized number. In addition, the prescription and dispensing are linked at prescription level. Moreover, in the birth registry there is a linkage between the mother and the newborn through their ID numbers

## Linked data sources

### **Pre linked**

Is the data source described created by the linkage of other data sources?

Yes

---

**Data source, other**

All data sources that composes VID (01\_SIP, 02\_PCV, 03\_CEX, 04\_MBDS, 05\_AED, 06\_DIAGNOSES, 07\_GAIA, 08\_SIV, 09\_MDR, 10\_PMR, 11\_EOS, 12\_TESTS, 13\_CONG and 14\_REDMIVA) have a linking ID number that identifies uniquely each person

---

**Linkage strategy**

Deterministic

---

**Linkage variable**

ID number (sip) for all the bases. Prescription number (receta\_id) for the prescription and dispensing linkage

---

**Linkage completeness**

The completeness of id number is 100% for each database, except for the birth registry, where the ID number of the newborn is always completed but the ID number of the mother is missing in a 18% of the cases.

## Data management specifications that apply for the data source

**Data source refresh**

Monthly

---

**Informed consent for use of data for research**

Required for intervention studies

---

**Possibility of data validation**

Can validity of the data in the data source be verified (e.g., access to original medical charts)?

Yes

---

### **Data source preservation**

Are records preserved in the data source indefinitely?

Yes

---

### **Data source preservation length (years)**

- years

---

### **Approval for publication**

Is an approval needed for publishing the results of a study using the data source?

Yes

---

### **Data source last refresh**

01/10/2025

## **Common Data Model (CDM) mapping**

### **CDM mapping**

Has the data source been converted (ETL-ed) to a common data model?

Yes

### **CDM Mappings**

#### **CDM name**

ConcepTION CDM

---

#### **CDM website**

<https://www.imi-conception.eu/>

---

**CDM release frequency**

6 months

---

**Data source ETL CDM version**

2.2

---

**Data source ETL frequency**

12,00 months

---

**Data source ETL specifications (file)**

[0\\_3\\_VID\\_Catalogue\\_RTL\\_specifications\\_0.pdf](#) (4.04 MB)

---

**Data source ETL status**

Completed

---

**Data source ETL specifications (link)**

[https://catalogues.ema.europa.eu/sites/default/files/document\\_files/0\\_3\\_VID\\_Cat...](https://catalogues.ema.europa.eu/sites/default/files/document_files/0_3_VID_Cat...)

---

**CDM name**

OMOP

---

**CDM website**

<https://www.ohdsi.org/Data-standardization/>

---

**Data source ETL CDM version**

5.4

---

**Data source ETL frequency**

12,00 months

---

**Data source ETL specifications (file)**

[FISABIO-HSRP OMOP ETL design v3.1.pdf](#) (3.18 MB)

---

**Data source ETL status**

In progress

---

**Data source ETL specifications (link)**

[https://catalogues.ema.europa.eu/sites/default/files/document\\_files/1\\_1\\_FISABIO...](https://catalogues.ema.europa.eu/sites/default/files/document_files/1_1_FISABIO...)