The Valencia Health System Integrated Database

First published: 01/02/2024

Last updated: 23/04/2024

Data source   Hospital outpatient visit records   Other   Primary care medical records
Registration with healthcare system

Administrative details

Administrative details

PURI
https://redirect.ema.europa.eu/resource/1111174

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

Data source website
https://www.san.gva.es/es/web/investigacio

Contact details

Francisco Sánchez-Sáez

(**Main**)

francisco.sanchez@fisabio.es

Gabriel Sanfélix-Gimeno

(**Alternate**)

gabriel.sanfelix@fisabio.es

Data source regions and languages

Data source countries
Spain

Data source languages
Catalan
English
Spanish

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

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)

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

SAFETY-VAC: a framework for the post-authorisation safety monitoring and evaluation of vaccines in the European Union (SAFETY-VAC)

The utilisation of antiepileptics in men and women of childbearing age, and pregnant women in Europe (ADEPT)

SAFETY-VAC: a framework for the postauthorisation safety monitoring and evaluation of vaccines in the European Union (SAFETY-VAC)

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
ICD-9-CM
Other

Prescriptions of medicines

Captured

Prescriptions vocabulary

ATC

Dispensing of medicines

Captured

Dispensing vocabulary

ATC

Advance 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

Quality of life measurements
Not Captured

Lifestyle factors
Captured

Lifestyle factors
Alcohol use
Tobacco use

Sociodemographic information
Captured

Sociodemographic information collected
Age
Country of origin
Deprivation index
Gender
Health area
Socioeconomic status

Quantitative descriptors

Population Qualitative Data

Population age groups
Paediatric Population (< 18 years)
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)
Elderly (? 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 between the following persons
Mother-child

Population

Population size
5600000

Active population

Active population size
5000000

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

Access and validation

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.gva.es/es/inicio/procedimientos?id_proc=14962

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

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

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

Data source last refresh
01/04/2023

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
6,00 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…

CDM name
OMOP

CDM website
https://www.ohdsi.org/Data-standardization/

Data source ETL CDM version
5.4

Data source ETL frequency
6,00 months

Data source ETL specifications (file)
1_1_FISABIO_HSRP_ETL_Design.pdf(6.62 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…