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

First published: 01/02/2024

Last updated: 12/11/2024

Data source | Human | Hospital discharge records | Pharmacy dispensing records | Primary care medical records

Administrative details

Administrative details

PURI

https://redirect.ema.europa.eu/resource/21501

Data source ID

21501

Data source acronym

BIFAP

Data holder

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

Data source type

Hospital discharge records

Pharmacy dispensing records

Primary care medical records

Main financial support

Funding by own institution

Care setting

Hospital inpatient care

Primary care – GP, community pharmacist level

Primary care – specialist level (e.g. paediatricians)

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

http://www.bifap.org/index_EN.html

Contact details

Miguel-Angel Macia

Main

mmacia@aemps.es

Edurne Lazaro

Alternate

elazaro@aemps.es

Data source regions and languages

Data source countries

Spain

Data source languages

Spanish

Data source regions

Aragón

Asturias, Principado de

Canarias

Cantabria

Castilla y León

Castilla-La Mancha

Madrid, Comunidad de

Murcia, Región de

Navarra, Comunidad Foral de

Data source establishment

Data source established

15/06/2001

Data source time span

First collection: 01/01/2002

The date when data started to be collected or extracted.

Publications

Data source publications

Updated list of Scientific Publications in peer reviewed journals describing the use of BIFAP for pharmacoepidemiology research linked to their respective DOIs.

Maciá-Martínez M-A, Gil M, Huerta C, et al. Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria (BIFAP): A data resource for pharmacoepidemiology in Spain. Pharmacoepidemiol Drug Saf. 2020;1–10. https://doi.org/10.1002/pds.5006

Studies

List of studies that have been conducted using the data source

ADVANCE POC I Risk pillar - Testing new approaches to monitoring benefit/risk with pertussis vaccines as test case: Incidence rates of safety outcomes of whole-cell pertussis and acellular pertussis vaccines in pre-school children

ADVANCE POC Study Protocol - Testing new approaches to monitoring benefit/risk with pertussis vaccines as test case. Coverage rates of acellular and whole-cell pertussis-containing vaccines in preschool children (ADVANCE Coverage POC)

Testing new approaches to monitoring benefit/risk with pertussis vaccines as test case: Incidence rates of pertussis and pertussis related outcomes of whole-cell pertussis and acellular pertussis vaccines in pre-school children (benefit

study on pertussis vaccination)

Exposure and coverage to routine schedule vaccines in different EU countries (ADVANCE-POC2)

The risk of acute liver injury associated with the use of antibiotics. A methodological comparison across epidemiological data sources

NESTED CASE-CONTROL STUDY TO ASSESS THE ASSOCIATION BETWEEN THE USE OF METHYLPHENIDATE AND THE RISK OF VALVULAR HEART DISEASE AND PULMONARY HYPERTENSION

USE OF DRUGS ACTING ON RENIN-ANGIOTENSIN SYSTEM (RAS) AND RISK OF COVID-19: A CASE-POPULATION STUDY (SRAA-COVID19)

Real-world effectiveness of different COVID-19 vaccines in Spain: a cohort study based on public electronic health records (BIFAP) (effectiveness of COVID-19 vaccines in Spain)

Characterising the risk of major bleeding in patients with Non-Valvular Atrial Fibrillation: non-interventional study of patients taking Direct Oral Anticoagulants in the EU

Metamizole and risk of agranulocytosis

Risk of thromboembolic events and thrombocytopenia after vaccination against COVID-19 (Thrombosis risk COVID-19 vaccination)

Establish an EU catalogue of sources of real-world data, characterised by a common set of metadata and data quality measurements

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

Impact of EU label changes and revised pregnancy prevention programme for oral retinoid containing medicinal products: utilization and prescribing trends

Impact of EU label changes and revised pregnancy prevention programme for medicinal products containing valproate: utilisation and prescribing trends

Effectiveness of heterologous and booster Covid-19 vaccination in 5 European countries, using a cohort approach in children and adults with a full primary Covid-19 vaccination regimen (Covid Vaccines Effectiveness (CoVE))

Cohort monitoring of Adverse Events of Special Interest and COVID-19 diagnoses prior to and after COVID-19 vaccination (ECVM)

Rapid Safety Assessment of SARS-CoV-2 vaccines in EU Member States using electronic health care datasources (CVM Covid19-Vaccine-Monitor-EHR)

Background rates of Adverse Events of Special Interest for monitoring COVID-19 vaccines (ACCESS-BGR)

Effectiveness of antiresorptives in preventing hip fractures in older women (≥ 75 years) with osteoporosis: nested case-control study cohort (BiHip)

DARWIN EU® - Chondrosarcoma: patient demographics, treatments, and survival in the period 2010-2023

DARWIN EU® - Characterising interstitial lung disease in Europe

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

DARWIN EU® – Trends in utilisation of Attention-Deficit Hyperactivity Disorder (ADHD) Medications

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

Detection of therapeutic cascades associated with gabapentinoids and benzodiazepines in adults over 65 years old in Spain using BIFAP through prescription sequence symmetry analysis (GABALOOP)

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

INCIDENCIA Y PREVALENCIA DE HIPOTIROIDISMO EN ESPAÑA, COMORBILIDAD, TRATAMIENTO Y ASOCIACION CON EFECTOS ADVERSOS DE SALUD (GRACHIPES)

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

DARWIN EU® – Paracetamol prescribing and paracetamol overdose in Europe: a descriptive analysis of trends and patient characteristics

Asymptomatic hyperuricemia: to treat or not to treat. A target trial emulation to assess major cardiorenal outcomes (HYPER-TTE-HARV)

DARWIN EU® - Suicidality incidence rates in adult male patients and in patients treated with finasteride and dutasteride

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.

No

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?

No

Cause of death

Captured

Cause of death vocabulary

ICD-10-CM

ICD-9-CM

Prescriptions of medicines

Captured

Prescriptions vocabulary

ATC

other

Prescriptions vocabulary, other

SNOMED

Dispensing of medicines

Captured

Dispensing vocabulary

ATC

other

Dispensing vocabulary, other

SNOMED

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

SNOMED CT

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-CM

ICD-9-CM

SNOMED CT

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?

No

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

Not coded (Free text)

SNOMED CT

Medicinal product information

Captured

Medicinal product information collected

Brand name

Dosage regime

Formulation

Package size

Strength

Medicinal product vocabulary

SNOMED

Quality of life measurements

Not Captured

Lifestyle factors

Captured

Lifestyle factors

Alcohol use

Other

Tobacco use

Sociodemographic information

Captured

Sociodemographic information collected

Age

Gender

Sex

Quantitative descriptors

Population Qualitative Data

Population age groups

All

Paediatric Population (< 18 years)

Neonate

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

92% of population in nine Spanish regions

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)

People who are registered with a primary care physician within the Spanish NHS (=98,9% of the Spanish population) in the 9 out of the 17 Spanish regions contribute data. Regional Healthcare Services act as health service providers for all levels of care in their territories. Only a proportion of civil servants (around 4% of the Spanish population) opts out of the general public system and they are covered by the private sector. There is a 13% of the Spanish population that contracts with private-for-profit health insurance companies, with an important regional variation. However, in general private health schemes do not cover outpatient prescription medicines and these patients usually are also registered and attend NHS for healthcare and reimbursement. This yields to only 1.1% of the Spanish population are not registered with a primary care physician within the Spanish NHS.

Population

Population size

22580036

Active population size

17384174

Population by age group

Age group	Population size	Active population size
Paediatric Population (< 18 years)	3174475	2664591
Neonate	105709	81731
Infants and toddlers (28 days – 23 months)	147212	116259
Children (2 to < 12 years)	1703964	1410787
Adolescents (12 to < 18 years)	1217590	1055814
Adults (18 to < 46 years)	7608667	5930964
Adults (46 to < 65 years)	5923021	5089694
Elderly (≥ 65 years)	5873873	3698925
Adults (65 to < 75 years)	2269983	1784108
Adults (75 to < 85 years)	1913469	1215680
Adults (85 years and over)	1690421	699137

Median observation time

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

10.00

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

12.00

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).

BIFAP Data Access Governance

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

Please see Section 3.1 'Data Collection' in GOVERNANCE DOCUMENT in: http://www.bifap.org/data-governance?lang=en

Event triggering registration

Event triggering registration of a person in the data source

Other

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

Upon registration with a primary care physician within the Spanish NHS (=98,9% of the Spanish population) in the 9 out of the 17 Spanish regions that contribute data

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

Death

Emigration

Event triggering creation of a record in the data source

In every encounter with the general practitioner/paediatrician. Hospital admission and pharmacy dispensation will also trigger the creation of a record.

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

Linked data sources

Pre linked

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

No

Data source, other

BIFAP Diagnosis Tests of Covid-19

Linkage strategy

Deterministic

Linkage variable

The Personal Identification Code for the Autonomous Community (CIPA) is the variable used in the Autonomous Communities for the combination or linking of the different records for the same patient (primary care records, pharmacy dispensing records, hospital discharge diagnoses and Other Data banks). Subsequently a pseudonymisation procedure is carried out: the CIPA is eliminated by the computer technicians of the autonomous community, maintaining a pseudonymised identifier that distinguishes each patient from the Others in BIFAP. At the AEMPS, the Patient ID is finally generated from the first identifier.

Linkage completeness

Only patients included the Primary care medical records (PC-MR) and identified with a CIPA (the single linkage variable) are linked to the Other Data sources in BIFAP. Therefore, completeness is 100%. On the Other hand, any patients identified with a CIPA in Data sources (Data banks) but not having a PC-MR are not included in BIFAP.

Pre linked

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

No

Data source, other

BIFAP EMRs from Primary Care

Linkage strategy

Deterministic

Linkage variable

The Personal Identification Code for the Autonomous Community (CIPA) is the variable used in the Autonomous Communities for the combination or linking of the different records for the same patient (primary care records, pharmacy dispensing records, hospital discharge diagnoses and Other Data banks). Subsequently a pseudonymisation procedure is carried out: the CIPA is eliminated by the computer technicians of the autonomous community, maintaining a pseudonymised identifier that distinguishes each patient from the Others in BIFAP. At the AEMPS, the Patient ID is finally generated from the first identifier.

Linkage completeness

Only patients included the Primary care medical records (PC-MR) and identified with a CIPA (the single linkage variable) are linked to the Other Data sources in BIFAP. Therefore, completeness is 100%. On the Other hand, any patients identified with a CIPA in Data sources (Data banks) but not having a PC-MR are not included in BIFAP.

Pre linked

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

No

Data source, other

BIFAP Hospital Diagnosis at in patients discharge

Linkage strategy

Deterministic

Linkage variable

The Personal Identification Code for the Autonomous Community (CIPA) is the variable used in the Autonomous Communities for the combination or linking of the different records for the same patient (primary care records, pharmacy dispensing records, hospital discharge diagnoses and Other Data banks). Subsequently a pseudonymisation procedure is carried out: the CIPA is eliminated by the computer technicians of the autonomous community, maintaining a pseudonymised identifier that distinguishes each patient from the Others in BIFAP. At the AEMPS, the Patient ID is finally generated from the first identifier.

Linkage completeness

Only patients included the Primary care medical records (PC-MR) and identified with a CIPA (the single linkage variable) are linked to the Other Data sources in BIFAP. Therefore, completeness is 100%. On the Other hand, any patients identified with a CIPA in Data sources (Data banks) but not having a PC-MR are not included in BIFAP.

Pre linked

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

No

Data source, other

BIFAP Medicines Dispensed at Community Pharmacies

Linkage strategy

Deterministic

Linkage variable

The Personal Identification Code for the Autonomous Community (CIPA) is the variable used in the Autonomous Communities for the combination or linking of the different records for the same patient (primary care records, pharmacy dispensing records, hospital discharge diagnoses and Other Data banks). Subsequently a pseudonymisation procedure is carried out: the CIPA is eliminated by the computer technicians of the autonomous community, maintaining a pseudonymised identifier that distinguishes each patient from the Others in BIFAP. At the AEMPS, the Patient ID is finally generated from the first identifier.

Linkage completeness

Only patients included the Primary care medical records (PC-MR) and identified with a CIPA (the single linkage variable) are linked to the Other Data sources in BIFAP. Therefore, completeness is 100%. On the Other hand, any patients identified with a CIPA in Data sources (Data banks) but not having a PC-MR are not included in BIFAP.

Pre linked

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

No

Data source, other

BIFAP Vaccines Covid-19 administered National Registry

Linkage strategy

Deterministic

Linkage variable

The Personal Identification Code for the Autonomous Community (CIPA) is the variable used in the Autonomous Communities for the combination or linking of the different records for the same patient (primary care records, pharmacy dispensing records, hospital discharge diagnoses and Other Data banks). Subsequently a pseudonymisation procedure is carried out: the CIPA is eliminated by the computer technicians of the autonomous community, maintaining a pseudonymised identifier that distinguishes each patient from the Others in BIFAP. At the AEMPS, the Patient ID is finally generated from the first identifier.

Linkage completeness

Only patients included the Primary care medical records (PC-MR) and identified with a CIPA (the single linkage variable) are linked to the Other Data sources in BIFAP. Therefore, completeness is 100%. On the Other hand, any patients identified with a CIPA in Data sources (Data banks) but not having a PC-MR are not included in BIFAP.

Pre linked

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

No

Data source, other

Causes of Death national registry

Linkage strategy

Deterministic

Linkage variable

The Personal Identification Code for the Autonomous Community (CIPA) is the variable used in the Autonomous Communities for the combination or linking of the different records for the same patient (primary care records, pharmacy dispensing records, hospital discharge diagnoses and Other Data banks). Subsequently a pseudonymisation procedure is carried out: the CIPA is eliminated by the computer technicians of the autonomous community, maintaining a pseudonymised identifier that distinguishes each patient from the Others in BIFAP. At the AEMPS, the Patient ID is finally generated from the first identifier.

In the Causes of Death national registry persons personal identifiers are different from the CIPA. Therefore, an additional linkage process of the CIPA and the personal identifiers in the Causes of Death national registry (national identitiy number and Others) is performed by the autonomous regions.

Linkage completeness

Only patients included the Primary care medical records (PC-MR) and identified with a CIPA (the single linkage variable) are linked to the Other Data sources in BIFAP. Therefore, completeness is 100%. On the Other hand, any patients with a record in the Cause of Death national registry but not having a PC-MR are not included in BIFAP.

Pre linked

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

No

Data source, other

Hospital Pharmacies dispensing Data

Linkage strategy

Deterministic

Linkage variable

The Personal Identification Code for the Autonomous Community (CIPA) is the variable used in the Autonomous Communities for the combination or linking of the different records for the same patient (primary care records, pharmacy dispensing records, hospital discharge diagnoses and Other Data banks). Subsequently a pseudonymisation procedure is carried out: the CIPA is eliminated by the computer technicians of the autonomous community, maintaining a pseudonymised identifier that distinguishes each patient from the Others in BIFAP. At the AEMPS, the Patient ID is finally generated from the first identifier.

Linkage completeness

Only patients included the Primary care medical records (PC-MR) and identified with a CIPA (the single linkage variable) are linked to the Other Data sources in BIFAP. Therefore, completeness is 100%. On the Other hand, any patients identified with a CIPA in Data sources (Data banks) but not having a PC-MR are not included in BIFAP.

Data management specifications that apply for the data source

Data source refresh

Every 6 months

Informed consent for use of data for research

Not Required

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?

No

Data source last refresh

31/12/2023

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

BIFAP

CDM website

http://www.bifap.org/docs/BIFAP_Data_Access_Governance_v2_2021_20211011_with%20.

CDM release frequency

6 months

Data source ETL CDM version

1.0

Data source ETL frequency

2,00 months

Data source ETL status

Completed

CDM name

ConcepTION CDM

CDM website

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

CDM release frequency

6 months

Data source ETL CDM version

For ConcepTION each study implies one or more ETL and different versions of the CDM can be used in different moments.

Data source ETL frequency

6,00 months

Data source ETL status

Completed

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 status

Completed