# IATROS

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

**Last updated:** 24/07/2025

Data source | Human | Hospital inpatient records | Hospital outpatient visit records | Other

## Administrative details

## Administrative details

### Data source ID

1111141

### Data source acronym

APHM-IATROS

### Data holder

Assitance Publique des Hopitaux de Marseille (APHM)

### Data source type

Hospital inpatient records

Hospital outpatient visit records

Other

**Data source type, other**

Electronic health records,Pharmacy dispensation 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


## Antoine Aka antoine.aka@ap-hm.fr

**Main**

antoine.aka@ap-hm.fr


# Data source regions and languages

**Data source countries**

France

**Data source languages**

French

**Data source regions**

Provence-Alpes-Côte-d'Azur

## Data source establishment

**Data source established**

01/01/2014

**Data source time span**

**First collection:** 01/01/2014

The date when data started to be collected or extracted.

# Publications

## Data source publications

Contextualising adverse events of special interest to characterise the baseline incidence rates in 24 million patients with COVID-19 across 26 databases: a multinational retrospective cohort study

Increased in-hospital mortality from COVID-19 in patients with schizophrenia.

Beta-lactam allergy labeling in intensive care units: An observational, retrospective study.

# 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

## 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

## Prescriptions of medicines

Captured

### Prescriptions vocabulary

ATC

other

### Prescriptions vocabulary, other

UCD (Common Dispensing Unit)

### Dispensing of medicines

Not Captured

### Dispensing vocabulary, other

UCD (Common Dispensing Unit)

### 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?

No

### Indication for use

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

Not Captured

### Medical devices

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

Yes

## Administration of vaccines

No

## Procedures

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

Captured

## Procedures vocabulary

Other

## Procedures vocabulary, other

CCAM (Common Classification of Aedical Acts)

## 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.

Captured

---

### Biomarker data vocabulary

Other

---

### Biomarker vocabulary, other

Our own vocabulary for now. We migrate to loinc in a few months

---

### Patient-reported outcomes

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

Yes

---

### 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

---

**Medicinal product information**

Captured

---

**Medicinal product information collected**

Active ingredient(s)

Brand name

Dose

Route of administration

---

**Medicinal product vocabulary**

Other

---

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

UCD (Common Dispensing Unit)

---

**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

Marital status

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**

40%

---

**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)**

No data on the private sector or primary care (this includes both visits to primary care and pharmacists). We only have data pertaining to hospital-based care (outpatient visits, inpatient hospitalisations, prescriptions).

## 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**

Mother-child

## Population

**Population size**

2329771

---

**Active population size**

17669

## Population by age group

| Age group | Population size | Active population size |
|---|---|---|
| Paediatric Population (< 18 years) | 450071 | 2612 |
| Preterm newborn infants (0 – 27 days) | 8638 | 24 |
| Term newborn infants (0 – 27 days) | 20551 | 325 |
| Infants and toddlers (28 days – 23 months) | 60166 | 475 |
| Children (2 to < 12 years) | 223642 | 1124 |
| Adolescents (12 to < 18 years) | 95961 | 761 |
| Adults (18 to < 46 years) | 817767 | 5839 |
| Adults (46 to < 65 years) | 481321 | 4345 |
| Elderly (≥ 65 years) | 467062 | 4873 |
| Adults (65 to < 75 years) | 223541 | 2382 |
| Adults (75 to < 85 years) | 156271 | 1767 |
| Adults (85 years and over) | 87250 | 533 |

## Median observation time

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

1.00

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

3.00

# Data flows and management

## Access and validation

### 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

The data collection process starts with a business and technical scoping. Data source experts and medical information department experts identify the variables that meet our use cases.

An ELT was developed in JAVA to build a datalake. Several data cleaning (data quality) and data transformation operations are then performed to obtain our database. The DBMS used is postgresql.

Several tests are performed on this ELT:
Data integration test - confirms that data from all sources has been loaded correctly into the target datalake, with threshold values checked.

Source-target data test - Ensures that the intended data is injected into the target system without being lost or truncated, and also that the number of records loaded into the datalake corresponds to the different sources.

## Event triggering registration

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

not availabe.

The source database is updated by Cron jobs, which are scheduled to run overnight.

## 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**

Air pollution index

---

**Linkage strategy**

Deterministic

**Linkage variable**

Geographical address

**Pre linked**

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

Yes

**Data source, other**

AXIGATE

**Linkage strategy**

Deterministic

**Linkage variable**

IEP : All data sources have this variable. IEP uniquely identifies an admission to the APHM, it is associated with the IPP which is the patient's number.

An IPP can have one or more IEPs

**Pre linked**

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

Yes

**Data source, other**

CORA(PMSI)

**Linkage strategy**

Deterministic

---

**Linkage variable**

IEP : All data sources have this variable. IEP uniquely identifies an admission to the APHM, it is associated with the IPP which is the patient's number.

An IPP can have one or more IEPs

**Pre linked**

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

No

---

**Data source, other**

Deprivation Index

---

**Linkage strategy**

Deterministic

---

**Linkage variable**

Geographical address

**Pre linked**

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

No

---

**Data source, other**

Insee death

---

**Linkage strategy**

Combination

**Linkage variable**

Lastname, firstname,date of birth

**Pre linked**

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

Yes

**Data source, other**

PHARMA

**Linkage strategy**

Deterministic

**Linkage variable**

IEP : All data sources have this variable. IEP uniquely identifies an admission to the APHM, it is associated with the IPP which is the patient's number.

An IPP can have one or more IEPs

## Data management specifications that apply for the data source

**Data source refresh**

Monthly

### Informed consent for use of data for research

Required for all 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

01/01/2025

## 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

OMOP

**CDM website**

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

---

**Data source ETL CDM version**

1

---

**Data source ETL frequency**

4,00 months

---

**Data source ETL status**

Completed