

# InGef Research Database

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

**Last updated:** 19/02/2026

Data source

Human

Administrative healthcare records (e.g., claims)

## Administrative details

### Administrative details

**Data source ID**

1111207

---

**Data source acronym**

InGef RDB

---

**Data holder**

[Institute for Applied Health Research Berlin \(InGef\)](#)

---

**Data source type**

Administrative healthcare records (e.g., claims)

---

**Main financial support**

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.ingef.de>

## Contact details

Raeleesha Norris [raeleesha.norris@ingef.de](mailto:raeleesha.norris@ingef.de)

**Main**

[raeleesha.norris@ingef.de](mailto:raeleesha.norris@ingef.de)

Anne Rothhardt [anne.rothhardt@ingef.de](mailto:anne.rothhardt@ingef.de)

**Alternate**

[anne.rothhardt@ingef.de](mailto:anne.rothhardt@ingef.de)

## Data source regions and languages

### **Data source countries**

Germany

---

### **Data source languages**

German

## Data source establishment

### **Data source established**

01/11/2015

---

### **Data source time span**

**First collection:** 01/01/2016

The date when data started to be collected or extracted.

## Publications

### Data source publications

[Characteristics and external validity of the German Health Risk Institute \(HRI\) Database](#)

[Effectiveness and safety of direct oral anticoagulants with antiplatelet agents in patients with venous thromboembolism: A multi-database cohort study](#)

[Sodium-Glucose Co-Transporter 2 Inhibitors and the Risk of Venous Thromboembolism in Patients with Type 2 Diabetes: A Cohort Study](#)

[Clinical outcomes and characteristics of patients hospitalized for Influenza or COVID-19 in Germany](#)

[Sampling strategy, characteristics and representativeness of the InGef research database](#)

## Studies

List of studies that have been conducted using the data source

Real World Outcomes of Patients Treated with Vericiguat in German Routine Care (ROVER)

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

DARWIN EU® - Prevalence of hypertrophic cardiomyopathy (HCM) and obstructive hypertrophic cardiomyopathy (oHCM) in six European countries

DARWIN EU® - Monitoring prescription of medicines for public health emergencies at risk of shortages

DARWIN EU® - Use of antiretroviral therapies in paediatric patients

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

DARWIN EU® - Antipsychotic prescribing in children in Europe: a descriptive analysis of trends and patient characteristics

DARWIN EU® - RR2 Drug utilisation study of prescription opioids

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

DARWIN EU® - Descriptive study of tetanus immunoglobulin use and tetanus-prone wounds in Europe

DARWIN EU® - RR Childhood hypertension and sartans prescribing in children

DARWIN EU® - Prevalence of selected cancers

DARWIN EU® - Incidence, period prevalence, and characterisation of individuals with paediatric pulmonary arterial hypertension

DARWIN EU® - Multiple myeloma: patient characterisation, treatments, and survival in the period 2012–2024

DARWIN EU® - Acute myeloid leukaemia: incidence, patient characteristics, treatments, and survival in the period 2015–2024

DARWIN EU® - Enabling pregnancy and mother-child research in DARWIN EU® Data Network - PeriNet

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

DARWIN EU® - Alzheimer's Disease: Incidence, Prevalence, and Individuals' Characteristics

DARWIN EU® - Treatment characterisation and post-diagnosis outcomes in Alzheimer's disease

DARWIN EU® - Determinants for use of GLP1 receptor agonists – a drug utilisation study

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

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

DARWIN EU - Background incidence rates of selected vaccine adverse events of special interest (AESIs) in Europe – supporting international preparedness

## 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.

No

---

## **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**

Not Captured

---

## **Prescriptions of medicines**

Captured

---

## **Prescriptions vocabulary**

ATC

other

---

## **Prescriptions vocabulary, other**

OPS (Operationen- und Prozedurenschlüssel) for some medications administered at hospital

---

## **Dispensing of medicines**

Captured

---

## **Dispensing vocabulary**

ATC

---

## **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)].

Yes

---

## **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?

Not Captured

---

## **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**

Other

---

## **Procedures vocabulary, other**

EBM (Einheitlicher Bewertungsmaßstab - doctor's fee scale) for ambulatory procedures; OPS (Operationen- und Prozedurenschlüssel) for operations conducted at the hospital

---

## **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?

No

---

## **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**

Other

---

### **Diagnosis / medical event vocabulary, other**

ICD-10-GM

---

### **Medicinal product information**

Captured

---

### **Medicinal product information collected**

Active ingredient(s)

Brand name

Dose

Package size

Route of administration

---

### **Medicinal product vocabulary**

ATC

Other

---

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

PZN (Pharmazentralnummer - pharmaceutical reference number)

---

### **Quality of life measurements**

Not Captured

---

### **Lifestyle factors**

Not Captured

---

### **Sociodemographic information**

Captured

---

### **Sociodemographic information collected**

Age

Deprivation index

Education level

Gender

Living in rural area

## 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)

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**

Approx. 10.5 Million insureds are included in the database, 7.8 Million of these actively insured in 2025. This corresponds to 9% of the total German population.

(26/28) Our database does not contain a direct distinction between preterm and term newborns. Additionally, date of birth are only respresented on a quarterly level. We therefore counted all individuals under 2 years of age within the "infants and toddlers (28 days to 23 months)" section.

---

**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)**

Insurees of private health insurances as well as those insurees of other statutory health insurances which do not contribute to the InGef RDB

## Population

**Population size**

10672793

---

**Active population size**

7840010

## Population by age group

<b>Age group</b>	<b>Population size</b>	<b>Active population size</b>
Paediatric Population (< 18 years)	1762275	1319348
Infants and toddlers (28 days - 23 months)	204963	140912
Children (2 to < 12 years)	960249	719530

Age group	Population size	Active population size
Adolescents (12 to < 18 years)	597063	458906
Adult and elderly population ( $\geq 18$ years)	8910518	6520662
Adults (18 to < 65 years)	6995858	5115378
Adults (18 to < 46 years)	4126791	2776437
Adults (46 to < 65 years)	2869067	2338941
Elderly ( $\geq 65$ years)	1914660	1405284
Adults (65 to < 75 years)	949582	805679
Adults (75 to < 85 years)	601974	417550
Adults (85 years and over)	363104	182055

## 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) captured**

10.00

## 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**

Data of the InGef RDB is stored via Teradata SQL Assistant

# Event triggering registration

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

Insurance coverage start

---

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

Death

Insurance coverage end

---

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

Hospital admission, encounter of physicians (GPs, specialists), reimbursable receipts at the pharmacy

# 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?

No

# Data management specifications that apply for the data source

**Data source refresh**

Monthly

---

**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)?

No

---

**Data source preservation**

Are records preserved in the data source indefinitely?

No

---

**Data source preservation length (years)**

10 years

---

**Approval for publication**

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

Yes

---

**Data source last refresh**

02/02/2026

## 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**

5.4

---

**Data source ETL frequency**

6,00 months

---

**Data source ETL status**

Completed