SAFETY-VAC study

A framework for the post-authorization SAFETY monitoring and evaluation of the VACcines in Europe

Specific Contract No 02 implementing framework contract No EMA/2020/46/TDA/18

Study Report for Objective 1:
Network of Data Sources for Vaccine Safety Evaluation

Deliverable 2 – April 2024

Main Authors: Carlos E. Durán, Rosa Gini, Davide Messina, Fabio Riefolo, Judit Riera, Vjola Hoxhaj, Sima Mohammadi, Irene Pazos, Vera Ehrenstein, Giuseppe Roberto, Miriam Sturkenboom.

Document version: 1.0
First Release date: 19 April 2024
Second Release date: 29 May 2024
Final Release date: 05 July 2024

Disclaimer & acknowledgement

The research leading to these results was conducted as part of the activities of the EU PE&PV (Pharmacoepidemiology and Pharmacovigilance) Research Network (led by Utrecht University) in collaboration with the Vaccine Monitoring Collaboration for Europe network (VAC4EU).

The scientific work for this project was coordinated by the University Medical Center Utrecht. The project has received support from the European Medicines Agency under the Framework service contract nr EMA/2020/46/TDA/L5.06.

This report expresses the opinion of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties.

This study has been registered in the HMA-EMA Catalogue of Real-World Data under the EU PAS number: EUPAS1000000193.

Table of contents

1 TITLE

SAFETY-VAC: A framework for the post-authorisation SAFETY monitoring and evaluation of VACcines in Europe.

2 ABSTRACT

2.1 Title

SAFETY-VAC: a framework for the post-authorisation SAFETY monitoring and evaluation of VACcines in Europe.

2.2 Keywords

Vaccines, Safety, Real-world evidence, Post-Authorisation.

2.3 Rationale and background

Numerous vaccines based on novel technologies targeting different diseases are continuously under development and obtaining marketing authorization. However, safety assessment is often limited to pre-authorisation clinical trials and new concerns are expected to arise during the post-authorisation phase. Thus, it is essential to create, assess, and describe a network of real-world data sources that are fit-for-purpose to address upcoming safety related questions, and to do so in a timely manner.
The European Medicines Agency (EMA) together with the European Centre for Disease Prevention and Control (ECDC) established the Vaccine Monitoring Platform (VMP) and the objective of generating a real-world evidence (RWE) framework for post-authorisation safety evaluation that can be leveraged in case of a new public health emergency or a safety concern occurring with a novel, or a more characterised, vaccine authorized in the European Union (EU) and the European Economic Area (EEA).

2.4 Research questions and objectives

To provide and describe a network of real-world data sources for the evaluation of vaccine safety signals, and to assess its fitness-for-purpose in conducting vaccine safety studies.

1) To assess the data quality for the purpose of conducting safety studies in the network and to describe data source population, capture of routine immunizations, and selected outcomes.
2) To assess whether data are fit for purpose for conducting vaccine safety studies.

2.5 Study design

A multi-database cohort design study was conducted from January 1st, 2017, till the last data availability, specific for each data source. We accessed data from 10 different electronic health record data sources from the EU PE&PV and VAC4EU networks of seven EEA countries that have proven to be able to convert data (n=9) into the ConcePTION common data model (CDM) or are willing to do this (n=1).
Persons were included in the dynamic study population when they have (a) information on age and gender available, (b) at least one day of follow in the study period (1/1/2017- latest availability). Follow-up started at the latest date of any of the following dates: (a) day that one year of lookback is available during the study period, or at birth for those born during the study period. Follow-up finished at the earliest of the following dates: death, disenrollment, or

recommended end date. The recommended end date is the latest date that the Data Access Provider (DAP) recommends having information from data banks complete.

Outcomes

Thirty-nine events have been selected together with EMA to assess data sources preparedness for a wide range of outcomes. For these events, incidence and prevalence rates were calculated. Clinical definition forms and codes' lists including ICD-9, ICD-10, SNOMED, and ICPC codes have been generated using the standardized VAC4EU process for identifying the events. The events included are:

- Microangiopathy (MA)
- Acute coronary artery disease (CAD)
- Arrhythmia
- Myocarditis
- Pericarditis
- Venous Thromboembolism (VTE)
- Arterial thrombosis
- TTS (VTE, arterial thrombosis, or CVST with thrombocytopenia in 10 days)
- Pulmonary embolism (PE)
- Haemorrhagic stroke
- Disseminated intravascular coagulation (DIC)
- Cerebral venous sinus thrombosis (CVST)
- Generalised convulsion
- Guillain Barré Syndrome (GBS)
- Diabetes (type 1)
- Single organ cutaneous vasculitis (SOCV)
- Erythema multiforme (EM)
- Meningoencephalitis
- Acute disseminated encephalomyelitis (ADEM)
- Narcolepsy
- Thrombocytopenia (TP)
- Transverse myelitis
- Bells' palsy
- Kawasaki's disease (KD)
- Pancreatitis
- Rhabdomyolysis (RML)
- Severe cutaneous adverse reactions to drugs (SCARs)
- Sensorineural hearing loss (SNHL)
- Graves' disease (GD)
- Hashimoto's thyroiditis (HT)
- Auto-immune hepatitis (AIH)
- Polyarteritis nodosa (PAN)
- Rheumatoid arthritis (RA)
- Psoriatic arthropathies (PsA)
- Systemic lupus erythematosus (SLE)
- Idiopathic thrombocytopenic purpura (ITP)
- Erythema nodosum (EN)
- Multiple sclerosis
- Ulcerative colitis (UC)

Exposure
The following vaccines were included in the fit for purpose assessment and were assessed in specific cohorts, nested in the study cohort:

- Measles-containing vaccines (doses 1, 2)
- Diphteria, tetanus toxoid, and pertussis (dose 1, 2, 3)
- Haemophilus influenzae type B (doses 1, 2, 3)
- Hepatitis B (doses 1, 2, 3)
- Polio (doses 1, 2, 3)
- Pneumococcal conjugate vaccines (doses 1, 2)
- Varicella (dose 1)
- Bacille Calmette-Guérin vaccine (dose 1)
- Human papillomavirus vaccine (doses 1, 2)
- Rotavirus (doses 1, 2)
- Meningoccocal vaccine (doses 1, 2)
- Influenza vaccine (dose 1)
- COVID-19 vaccines (doses 1 to 6)

Covariates
The following covariates were assessed:

- Age
- Gender
- Transplantation
- Immunocompromised status
- Pregnancy
- Hypertension
- Lipid abnormalities
- Malignancies
- HIV
- Cardiocerebrovascular disease
- Heart failure
- Diabetes
- Valvular heart disease
- Inflammatory bowel disease
- Coronary artery disease
- Myocardial infarction
- Arrhythmia
- VTE
- Infection
- Liver disease
- Alcohol abuse
- Sepsis
- Chronic renal disease
- Dementia
- Respiratory infections
- Herpes simplex
- Influenza
- Sleep disorders
- Mental health diseases
- Preeclampsia

- Hepatitis C
- Rheumatoid arthritis
- SLE
- Dermatomyositis
- Sjogren's syndrome
- Gallstones
- Sickle cell disease
- Myasthenia gravis
- Pernicious anemia
- Autoimmune hepatitis
- Celiac disease
- Hepatitis B
- Psoriasis
- Gout
- Crohn's disease
- Ulcerative colitis
- Atopic dermatitis
- Immune thrombocytopenia
- Nonalcoholic fatty liver
- Obesity
- Dermatomyositis

Data Sources

The following data sources were included: CPRD from the United Kingdom (access provided by Utrecht University (UU)); BIFAP (access provided by the Spanish Agency for Medicines and Medical Devices (AEMPS)), SIDIAP (access provided by the Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAP JGol)), VID (access provided by the Foundation for the Promotion of Health and Biomedical Research of Valencia Region (FISABIO)), and EPICHRON (access provided by the public health system in Aragón (IACS)) from Spain; PEDIANET from Italy (access provided by Società Servizi Telematici (SoSeTe)); Danish National Registries (DHR) (access provided by Aarhus University); Norwegian registers (NHR) (access provided by University of Oslo); Finnish registers (access provided by University of Eastern Finland); and SNDS from France (access provided by Bordeaux PharmacoEpi platform).

Analysis

For assessing the data quality, INSIGHT data quality level 1 and 2 checks were required as well as the running of tailored scripts on the incidence and prevalence of selected events, the coverage of selected vaccines and the prevalence of covariates. The " Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework" (SPIFD) tool was used to assess and summarize the results. For the population we required that data were available from 1/1/2017 and that dates of birth are accurate. For childhood vaccines, we assessed whether birth cohorts 2019 and 2020 were related to specific WHO/ECDC coverage indicators. For other vaccines we used the literature. In general, more than 10% deviation meant that the source did not meet the requirement for that outcome. For events we used rates provided by data sources with both primary care, outpatient diagnoses and hospital diagnoses as the reference.

2.6 Results

Nine data sources had an available data instance (specific subset of information/data required for one or more studies that have been ETL'ed into the ConcePTION CDM and quality checked at a certain point in time by the data sources) in the ConcePTION CDM and passed the INSIGHT

level 1 (ETL verification) and level 2 (Logical checks) quality checks with the quality check auditor. Quality checks are generally applied to each data instance as soon as is produced. At this stage, level 3 checks were not run by all DAPs, as a tailored analysis script for this study was provided to estimate incidence and prevalence rates, as well as vaccine coverage, to be able to assess whether data sources would be fit for purpose for causal inference vaccine safety questions. Since there is not a specific research question yet, we provided a summary of the fitness of population, exposure, covariate and outcome data, based on the SPIFD tool. The summary is based on the currently available data instances. The population for which the descriptive script was run comprised a total of 53,283,613 persons.
Incidence rates of 39 events were created, and for 12 vaccine coverage indicators were produced and benchmarked against WHO and ECDC value indicators. Table A1 summarizes the assessment.

2.7 Conclusion

This study provides a rapid description of the content of 9 data instances from 6 countries that were available at the moment of the contract signature (February 15, 2024). Seven data sources from 5 countries produced the requested results (attrition, prevalence of covariates, coverage of selected vaccines, and incidence and prevalence rates of 39 outcomes). This fitness-for-purpose assessment was performed on available data instances based on datasets originally extracted and used for previous COVID-19 vaccine research. Therefore, information on other vaccines rather than COVID-19 as well as of the outcomes were incomplete. New dataset extractions with all vaccines are currently being done. As soon as these up-to-date data instances are available, a new fitness for purpose assessment will be conducted.
For a new study, the time to analyse and produce final results based on new data instances within the SAFETY-VAC study framework is 3 to 6 months depending on the data source and the specific design and data requirements. Three data sources could not provide the required data in time for this report due to slow ethical or scientific approvals (SNDS, EPICHRON, and Finnish registries). This shows the need to have requirements and data ready and converted into the CDM to address rapid questions for the Vaccine Monitoring Platform (VMP).

Figure A1. SPIFD heatmap assessment for current data instances

LEGEND		
	5	Many/nearly all data requirements met
	4	Several data requirements met
	3	Likely that several data requirements are met but require further investigation
	2	Some data requirements met or unable to assess at this time
	1	Data requirements not met

	Fast	Fast timelines (data access and analysis) < 3 months
	Moderate	Moderate timelines (e.g., to data access, to analyse) 4 to 6 months
	Slow	Slow timelines (e.g., to data access, to analyse) > 6 months

Study characteristics and considerations	Requested information	Data sources							
		BIFAP_PC-ES	BIFAP_HOSP-ES	SIDIAP-ES	VID-ES	PEDIANET-IT	NHR-NO	DHR-DK	CPRD-UK
DESIGN ELEMENTS									
Study population	• At-least one day of follow-up from 1/1/2017, plus one year look-back. • Age and gender information.	4	4	5	4	5	5	5	5
Vaccine exposure group	Measles-containing vaccines	4	5	5	1	4	5	1	1
	DTP	3	3	5	1	4	5	1	1
	Haemophilus influenzae type B	3	3	5	1	5	5	1	1
	Hepatitis B	3	3	5	1	4	5	1	1
	Polio	3	3	5	1	5	5	1	1
	Pneumococcal conjugate vaccines	5	5	3	1	3	3	1	1
	Varicella	5	5	1	1	5	1	1	1
	HPV	1	1	1	1	4	5	1	1
	Rotavirus	5	5	4	1	4	3	1	1
	Meningococcal vaccine	5	5	5	1	5	1	1	1
	Influenza vaccine	5	5	5	5	5	5	1	5
	COVID-19 vaccines	5	5	5	5	5	5	5	5
Primary outcomes (availability of events through	Acute coronary artery disease (CAD)	3	3	5	5	1	5	5	4
	ADEM	2	5	5	5	1	5	5	1
	Arrhythmia	5	5	5	5	5	5	5	5
	Arterial thrombosis	3	5	5	5	1	5	5	3

Study characteristics and considerations	Requested information	Data sources							
		BIFAP_PC-ES	BIFAP_HOSP-ES	SIDIAP-ES	VID-ES	PEDIANET-IT	NHR-NO	DHR-DK	CPRD-UK
diagnosis codes and drug proxies)	Autoimmune hepatitis	3	3	3	3	1	3	1	3
	Bell's palsy	5	5	5	5	3	4	5	5
	Cerebral venous sinus thrombosis	2	4	5	5	3	3	3	2
	DIC	1	2	4	4	1	4	4	1
	Erythema multiforme	4	4	5	5	5	5	5	4
	Erythema nodosum	5	5	5	5	5	5	1	1
	Generalized convulsion	3	3	5	5	5	5	5	5
	Haemorrhagic stroke	4	4	5	5	5	5	5	4
	Diabetes type 1	4	4	4	4	5	4	4	4
	Bell's palsy	5	5	5	5	5	5		5
	Grave´s disease	4	4	3	3	4	4	1	1
	Guillain Barré Syndrome	2	5	5	5	5	5	5	2
	Haemorrhagic stroke	4	4	5	5	1	5	5	4
	Hashimoto's thyroiditis	4	4	5	5	4	5	5	4
	Idiopathic thrombocytopenic purpura	4	4	5	5	3	4	4	4
	Kawasaki's disease	4	4	4	4	5	4	4	4
	Meningoencephalitis	2	4	5	5	2	5	4	2
	Microangiopathy	1	3	5	5	2	5	5	1
	Multiple sclerosis	1	4	5	5	1	5	5	1
	Myocarditis	3	4	5	5	1	5	5	4
	Narcolepsy	5	4	5	5	1	5	5	4
	Pancreatitis, acute	3	4	2	2	1	5	5	3
	Pericarditis	4	4	5	5	3	5	5	4
	Polyarteritis nodosa	4	4	5	5	1	5	1	4
	Psoriatic arthropathies	3	3	5	5	1	5	5	1
	Pulmonary embolism	2	4	5	5	1	5	5	4
	Rhabdomyolysis	2	3	4	4	1	3	4	3
	Rheumatoid arthritis	4	4	5	4	3	5	4	4
	SCAR	5	5	5	5	1	5	5	5
	Sensorineural hearing loss	3	3	3	3	3	3	1	3
	Single organ cutaneous vasculitis	4	5	5	5	1	5	5	4
	Systemic lupus erythematosus	4	4	5	4	1	5	5	4
	Thrombocytopenia	3	3	5	5	4	4	3	3
	Transverse myelitis	3	5	5	5	1	5	5	5

Study characteristics and considerations	Requested information	Data sources							
		BIFAP_PC-ES	BIFAP_HOSP-ES	SIDIAP-ES	VID-ES	PEDIANET-IT	NHR-NO	DHR-DK	CPRD-UK
	TTS	2	3	4	4	1	4	5	5
	Ulcerative colitis	5	4	5	5	2	5	5	1
	Venous thromboembolism	4	4	5	5	2	5	5	4
Confounding variables	Availability of key covariates at start of follow-up according to the protocol requirements.	5	5	5	5	5	5	5	5
Key subgroups	Availability to produce a pregnancy cohort	2	2	5	2	1	5	5	2
DATA ACCESS CONSIDERATIONS									
Timeline 1	Time to analyze based on the current instance	Fast	Fast	Fast	Fast	Fast	Fast	Fast	Fast
Timeline 2	Time to analyze based on new data instance within the SAFETY-VAC study framework	Fast	Fast	Fast	Moderate	Fast	Moderate	Slow	Fast

3 MARKETING AUTHORISATION HOLDER

Not applicable (NA)

4 INVESTIGATORS

University Medical Center Utrecht (UMCU), Utrecht, The Netherlands	Dr. Carlos E. Durán, Prof. Dr. Miriam Sturkenboom, Drs. Vjola Hoxhaj, Drs. Judit Riera Drs. N Luxi
Universiteit Utrecht (UU), Utrecht, The Netherlands CPRD data	Prof. Dr. Olaf Klungel, Dr. Patrick Souverein, Dr. Helga Gardarsdottir Dr. Satu Johanna Siiskonen
VAC4EU	Drs. Sima Mohammadi
Teamit Institute	Dr. Fabio Riefolo, Dr. Irene Pazos
Agenzia Regionale di Sanitá Toscana (ARS) Data Tuscany region	Dr. Rosa Gini, Davide Messina, Dr. Giuseppe Roberto Anna Girardi Giulia Hyeraci
Società Servizi Telematici -Pedianet	Prof. Dr. Carlo Giaquinto, Dr. Elisa Barbieri, Luca Stona
Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAP JGol)	Dr. Felipe Villalobos, Dr. Martín Solórzano Carlo Alberto Bissacco
Instituto Aragonés de Ciencias de la Salud (IACS)	Dr. Antonio Gimeno, Dr. Beatriz Poblador, Dr. Mercedes Aza, Dr. Aida Moreno, Alejandro Santos
Department of Clinical Epidemiology, Aarhus University and Aarhus University Hospital	Prof. Vera Ehrenstein, Lise Skovgaard Svingel, Benjamin Randeris Johannesen
Bordeaux PharmacoEpi platform (BPE) & ADERA	Dr. Cécile Droz-Perroteau, Laure Carcaillon-Bentata,
University of Eastern Finland	Prof. Anna-Maija Tolppanen, Prof. Sirpa Hartikainen, Dr. Thuan Vo, Dr. Anne Paakinaho, Blair Rajamaki
University of Oslo (UiO), Norway Norwegian linked registry data	Prof Dr. Hedvig ME.Nordeng, Saeed Hayati, Dr. Mahmoud Zidan
Foundation for the Promotion of Health and Biomedical Research of Valencia Region (FISABIO) – Valencia health system Integrated Database (VID)	Dr. Juan José Carreras Martínez, Dr. Arantxa Urchueguía Fornes, Elisa Correcher Martínez, Dr. Javier Díez-Domingo

Spanish Agency on Medicines and Medical Devices (AEMPS) -BIFAP database	Dr. Mar Martin, Dr. Patricia Garcia-Poza, Dr. Airam de Burgos, Belén Castillo-Cano, Dr. Elisa Martín-Merino

5 MILESTONES

Start of project	15 Feb 2024
D1 Project planning virtual meeting	28 Feb 2024
D2 Study report for Object 1	19 Apr 2024
D2 Study report for Object 1 acceptance	17 Jun 2024

6 RATIONALE AND BACKGROUND

The COVID-19 pandemic emphasised the public health need for comprehensive and rapid post-authorisation vaccine safety surveillance. An increasing number of vaccine products are based on novel technologies, for which safety experience is limited to pre-authorisation clinical trials until the recent COVID-19 pandemic. While new safety concerns are expected to arise with these novel vaccines, continuous monitoring and evaluation throughout the entire lifecycle remains necessary for authorized vaccines (1,2). To this aim, networks of real-world data sources that are fit-for-purpose and readily accessible are essential.

In May 2022, the EMA and the ECDC established the Vaccine Monitoring Platform (VMP)[1] with the perspective of generating RWE on the safety and effectiveness of vaccines in the EU and the European Economic Area (EEA) (3). The VMP research agenda, endorsed in July 2023 by the Immunisation and Vaccine Monitoring Advisory Board (IVMAB), confirmed the need for RWE capacity, capability and readiness to allow the timely evaluation of vaccine safety concerns. Therefore, EMA required the service of a contractor to provide a framework for post-authorisation safety evaluation that can be leveraged in case of a new public health emergency or a safety concern occurring with a novel or a more characterised vaccine.

7 RESEARCH AND OBJECTIVES

The overarching goal of the SAFETY-VAC project is to create a framework for the post-authorisation safety monitoring and evaluation of vaccines in Europe that can conduct near real-time studies on new or existing vaccines.

This report specifically addresses the **Objective 1** of the SAFETY-VAC project which aims to provide and describe a network of real-world data sources for the evaluation of vaccine safety signals, and to assess its fitness-for-purpose for regulatory studies through the following specific objectives:

[1] https://www.ema.europa.eu/en/about-us/what-we-do/crisis-preparedness-management/vaccine-monitoring-platform

Objective 1a. To assess the data quality for the purpose of conducting safety studies in the created network, including description of data source population, capture of routine vaccinations, and selected outcomes.

Objective 1b. To assess whether data are fit-for-purpose for conducting future safety studies on specific vaccines and selected outcomes in a near real-time monitoring manner.

8 AMENDMENTS AND UPDATES OF THE PROTOCOL

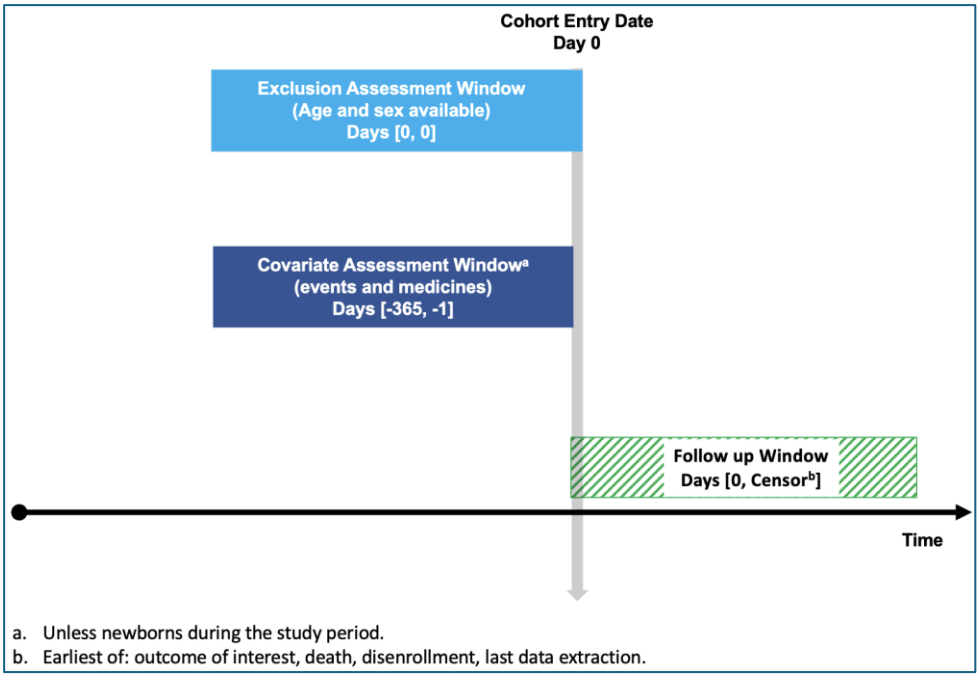
None

9 RESEARCH METHODS

9.1 Study design

We conducted a retrospective, multi-database, population-based cohort study, to describe and assess quality and the fitness for purpose of data sources to conduct potential future vaccine safety studies.

Figure 1 Study diagram for objective 1



9.2 Setting

In this study, a multi-database cohort has been constructed using electronic health records or registries data from 10 data sources in 7 European countries (Spain, Denmark, Finland, Norway, Italy, France, the United Kingdom) from January 1st, 2017, till the last data availability (see details in section 9.5 Data sources and measurement). The 10 different electronic health record data sources have proven to be able to convert data (n=9) into the ConcePTION common data model (CDM) or started the process (n=1).

9.3 Subjects

Persons were included in the study population when they had:

- Information on age and sex available.
- At least one day of follow up in the study period (1/1/2017- latest availability).
- At least one year lookback history, with the exception of newborns.

Follow up started on the latest of the following dates: study start date or date at which they have one year of lookback time, except for newborns during the study period, these were included upon birth or when born in the year before the study start. Follow-up finished at the earliest of the following dates: death, disenrollment, recommended end date by the DAP. The recommended end date is the date provided by the DAP until which they inform the data of different databanks is complete (see Figure 1). For calculation of incidence rates (IRs) of outcomes, the occurrence of the diagnosis of interest for the selected events was an additional censoring date.

9.4 Variables

Exposure

The Data Access Partners (DAPs) convert their local data on vaccinations into the VACCINES table of the CDM. Local vaccination information obtained from recorded prescription, dispensations, or administration by a general practitioner (GP) (see Table 1**),** were considered since they are part of the current vaccination schedule in one or more of the participating countries.

Table 1. Vaccines of interest for feasibility assessment

Indicator	Cohorts
Measles-contaning vaccine, dose 1	Childhood from birth to 24 months
Measles-contaning vaccine, dose 2	Childhood from birth to 24 months
Diphteria tetanus toxiod and pertussis, dose 1	Childhood from birth to 12 months
Diphteria tetanus toxiod and pertussis, dose 2	Childhood from birth to 12 months
Diphteria tetanus toxiod and pertussis, dose 3	Childhood from birth to 24 months
Haemophilus influenzae type B, dose 1	Childhood from birth to 12 months
Haemophilus influenzae type B, dose 2	Childhood from birth to 12 months
Haemophilus influenzae type B, dose 3	Childhood from birth to 24 months
Hepatatis B, dose 1	Childhood from birth to 12 months
Hepatatis B, dose 2	Childhood from birth to 12 months
Hepatitis B, dose 3	Childhood from birth to 24 months
Polio, dose 1	Childhood from birth to 12 months
Polio, dose 2	Childhood from birth to 12 months
Polio, dose 3	Childhood from birth to 24 months
Pneumococcal conjugate vaccines, dose 1	Childhood from birth to 12 months
Pneumococcal conjugate vaccines, dose 2	Childhood from birth to 24 months
Varicella	Childhood from birth to 24 months
Bacille Calmette-Guérin (BCG) vaccine, dose 1	Childhood from birth to 24 months
Human papillomavirus vaccine, dose 1	Adolescents, 9 to 15-year-old, stratified per gender
Human papillomavirus vaccine, dose 2	Adolescents, 9 to 15-year-old, stratified per gender
Rotavirus, dose 1	Childhood from birth to 12 months
Rotavirus, dose 2	Childhood from birth to 12 months
Meningococcal vaccine, dose 1	Childhood from birth to 12 months
Meningococcal vaccine, dose 2	Childhood from birth to 15 months

Indicator	Cohorts
Influenza vaccine	Seasonal cohorts, all ages, from 1 September to 30 April
COVID Vaccines, dose 1	Cohort entering 1 December 2020, stratified per age band at end of follow-up
COVID Vaccines, dose 2	Cohort entering 1 December 2020, stratified per age band at end of follow-up
COVID Vaccines, dose 3	Cohort entering 1 December 2020, stratified per age band at end of follow-up
COVID Vaccines, dose 4	Cohort entering 1 December 2020, stratified per age band at end of follow-up
COVID Vaccines, dose 5	Cohort entering 1 December 2020, stratified per age band at end of follow-up
COVID Vaccines, dose 6	Cohort entering 1 December 2020, stratified per age band at end of follow-up

The ATC codes or the vaccine type (vxtype) nomenclature were utilized by DAPs to store records of vaccination in their CDM instances during the ETL phase. The vaccine type is the only case when a semantic mapping is adopted during conversion to the ConcePTION CDM (4). The investigators retrieved from the level 1B checks (see **Annex 1**) the values utilised by the DAPs and mapped them to the corresponding code of antigens developed in the ADVANCE project (5). It is a code of three letters, and the vaccine is represented by the hyphen-separated alphabetic sequence of such three-letters codes. For example, all ATC codes of vaccine types associated to a trivalent diphteria-pertussis-tetanus vaccine is mapped to DIP-PER-TET.
In the study script, all vaccines corresponding to one of the three-digit codes were retrieved, using both vaccine type and ATC codes. Then, for each indicator of Table 6, all the corresponding vaccines were retrieved, possibly using the same record for multiple indicators. For example, a record of DIP-HIB-PER-POL-TET was replicated three times, once per the indicator DPT, once per the indicator HiB, and once for the indicator Pol. Vaccines that were less than 30 days apart were discarded, because they were considered duplicates. Finally, dose of vaccines was obtained from the sequence, independent of the recorded dose.

Outcomes:
Table 2 shows the outcomes that were selected in collaboration with EMA during the planning meeting and comprised: 1) AESI list for COVID-19 vaccines (ACCESS[2], SPEAC)[3], 2) AESI that might occur with vaccines in general, and 3) chronic immune mediated events that will be used for the other objectives in the project.

Table 2. List of selected events.

Num.	Name of the event
1	Microangiopathy (MA)
2	Acute coronary artery disease (CAD)
3	Arrhythmia
4	Myocarditis
5	Pericarditis
6	Venous thromboembolism (VTE)
7	Arterial thrombosis (AMI /Ischemic stroke)
8	TTS (VTE, arterial thrombosis, or CVST with thrombocytopenia in 10 days)
9	Pulmonary embolism (PE)

[2] https://zenodo.org/communities/vac4eu/records?q=ACCESS&l=list&p=1&s=10&sort=bestmatch
[3] https://zenodo.org/communities/speac_project/records?q=&l=list&p=1&s=10&sort=newest

Num.	Name of the event
10	Haemorrhagic stroke
11	Disseminated intravascular coagulation (DIC)
12	Cerebral venous sinus thrombosis (CVST)
13	Generalised convulsion
14	Guillain Barré Syndrome (GBS)
15	Diabetes (type 1)
16	Single organ cutaneous vasculitis (SOCV)
17	Erythema multiforme (EM)
18	Meningoencephalitis
19	Acute disseminated encephalomyelitis (ADEM)
20	Narcolepsy
21	Thrombocytopenia (TP)
22	Transverse myelitis
23	Bells' palsy
24	Kawasaki's disease (KD)
25	Pancreatitis
26	Rhabdomyolysis (RML)
27	Severe cutaneous adverse reactions to drugs (SCARs)
28	Sensorineural hearing loss (SNHL)
29	Graves' disease (GD)
30	Hashimoto's thyroiditis (HT)
31	Auto-immune hepatitis (AIH)
32	Polyarteritis nodosa (PAN)
33	Rheumatoid arthritis (RA)
34	Psoriatic arthropathies (PsA)
35	Systemic lupus erythematosus (SLE)
36	Idiopathic thrombocytopenic purpura (ITP)
37	Erythema nodosum (EN)
38	Multiple sclerosis
39	Ulcerative colitis (UC)

Event definition forms and code lists including ICD-9, ICD-10, SNOMED, and ICPC codes were used if available or generated using the following standard VAC4EU process. Event definition forms systematically capture the following items as a living document that will be closed and published upon study's end.

- Purpose of the event: covariate or outcome
- Version
- Document history
- Objective
- Clinical definition
- Synonyms/lay terms (for text mining purposes)
- Laboratory tests specific for diagnosing events
- Diagnostic tests specific for diagnosing events
- Drugs that are used to treat events
- Procedures used to treat events
- Setting where a condition is diagnosed (hospital, outpatient, GP)
- Literature review of diagnosis codes or algorithms used in other papers (health outcomes of interest)

- Code list
- Algorithm proposal
- References

Examples of completed event definition forms for the ACCESS study (6) can be found in the VAC4EU Zenodo repository (7).

Code lists

Code lists to identify outcomes were created using the VAC4EU Code Mapper tool (8), which maps concepts across vocabularies based on the Unified Medical Language System. Study variables are named in a standard VAC4EU hierarchical fashion based on the body system. The output of the Code Mapper is an Excel or CSV list. Each code is subsequently tagged as narrow or possible by two medical reviewers from the VAC4EU code list taskforce based on standard VAC4EU work instructions. Comments are consolidated in the VAC4EU code list task force.

The code lists are subsequently compiled in a CSV file through a standard R code which:

- Checks for ranges of codes in the Code Mapper outputs, and replacement with unique parent codes.
- Checks for odd characters in codes.
- Rounding of SNOMED codes.

Covariates

Table 3 describes the list of selected covariates. Covariates were selected based on risk factors of the events of interest, based on a literature review, if they existed. Code lists were created and tagged based on the same VAC4EU process as described under outcomes but now specific for covariates. Medication use may also be used as a proxy for comorbidities (see Table 3). Covariates were assessed in the cohort within a lookback of 365 days for diagnoses codes and for medicines.

Table 3. List of selected covariates and the CDM data that it is based on.

Covariate	**Source ConcePTION CDM tables**
Age	From PERSONS table
Gender	From PERSONS table
Race/ethnicity	From PERSONS table (if available)
Number of GP visits	From VISIT_OCCURRENCE table
Number of hospitalizations	From VISIT_OCCURRENCE table
Transplantation	From multiple tables: EVENTS, MEDICINES, SURVEY_OBSERVATIONS

Immunocompromised status	Algorithm from multiple tables:
	• EVENTS
	  ○ Inflammatory bowel disease
	  ○ Diabetes type 1
	  ○ Gout
	  ○ AIDS
	  ○ Sjogren syndrome
	  ○ Systemic Lupus Erythematosus
	  ○ Transplant recipient
	  ○ Psoriasis
	  ○ Psoriatic arthropathy
	  ○ Rheumatoid arthritis
	  ○ Spondylarthritis
	  ○ Multiple sclerosis
	  ○ Hematological cancer
	  ○ Multiple immunodeficiencies
	• MEDICINES
	  ○ Immunosuppressants

Pregnancy	From multiple tables (data source specific) using ConcePTION pregnancy algorithm (9)
Hypertension	EVENTS
Lipid abnormalities	EVENTS and MEDICINES
Malignancies	EVENTS and MEDICINES
HIV	EVENTS and MEDICINES
Decreased renal function	EVENTS
Cardiocerebrovascular disease	EVENTS and MEDICINES
Heart failure	EVENTS
Diabetes type II	EVENTS and MEDICINES
Valvular heart disease	EVENTS
Inflammatory bowel disease	EVENTS
Coronary artery disease	EVENTS
Myocardial infarction	EVENTS
Arrhythmia	EVENTS
VTE	EVENTS
Infection	EVENTS and MEDICINES
Liver disease	EVENTS
Alcohol abuse	EVENTS
Sepsis	EVENTS
Chronic renal disease	EVENTS
Dementia	EVENTS
Respiratory infections	EVENTS
Herpes simplex	EVENTS
Influenza	EVENTS
Sleep disorders	EVENTS
Mental health diseases	EVENTS and MEDICINES
Preeclampsia	EVENTS
Hepatitis C	EVENTS
Rheumatoid arthritis	EVENTS

Covariate	Source ConcePTION CDM tables
SLE	EVENTS
Dermatomyositis	EVENTS
Sjogren's syndrome EVENTS Gallstones	EVENTS
Sickle cell disease	EVENTS and MEDICINES
Myasthenia gravis	EVENTS
Pernicious anemia	EVENTS
Autoimmune hepatitis	EVENTS
Celiac disease	EVENTS
Hepatitis B	EVENTS
Psoriasis	EVENTS
Gout	EVENTS
Crohn's disease	EVENTS
Ulcerative colitis	EVENTS
Atopic dermatitis	EVENTS
Immune thrombocytopenia	EVENTS
Nonalcoholic fatty liver	EVENTS
Obesity	EVENTS and MEDICINES
Dermatomyositis	EVENTS

9.5 Data sources and measurement

This study used data from secondary electronic health record databases that are population-based. The characteristics of each of the participating DAPs are summarized in Table 4 and further detailed below.

Table 4. Data provider and data sources.

Country	Data Source	Data Access Provider	Estimated source population size	Start and end date of data instance*
Spain (ES)	BIFAP	BIFAP	17 million	1.1.2018-30.4.2022
Spain (ES)	SIDIAP	IDIAP JGol	5.8 million	1.1.2017-30.06.2023
Spain (ES)	VID	FISABIO	5.0 million	1.1.2018- 31.12.2022
Spain (ES)	EPICHRON	IACS	1.3 million	
Italy (IT)	PEDIANET	So.Se.Te	50.000	1.1.2011 (except hospitalizations, 1.1.2017) - 31.12.2022
Denmark (DK)	Danish national registries (DHR)	Aarhus University	5.9 million	1.1.2015- 31.12.2022
Norway (NO)	Norwegian national registers	University of Oslo	5.3 million	1.1.2017- 31.12.2022
United Kingdom (UK)	CPRD	Utrecht University	16 million	01.01.2017 - 31.12.2022
France (FR)	SNDS	BPE & ADERA	6.7 million (10% sample of the total population)	01.01.2017 – 31.12.2020
Finland (FI)	Finnish national registers	University of Eastern Finland	2.9 million (50% random sample of total population)	Not Available (n.a.)

* data instance is the subset of the data source that has been ETL'ed into the CDM at a certain point in time, this instance does not necessarily contain data from all databanks in the data source, but data required for one or more studies. Data quality checks are done for each data instance.

9.5.1 ES: BIFAP (SEVERAL REGIONS)

BIFAP (Base de Datos para la Investigación Farmacoepidemiológica en el Ámbito Público), a computerized database of medical records of primary care is a non-profit research project funded by the Spanish Agency for Medicines and Medical Devices (AEMPS). Information collected by PCPs includes administrative, socio-demographic, lifestyle, and other general data, clinical diagnosis and health problems, results of diagnostic procedures, interventions, and prescriptions/dispensations. Diagnoses are classified according to the International Classification of Primary Care (ICPC)-2, ICD-9 and SNOMEDCT system, and a variable proportion of clinical information is registered in "medical notes" in free text fields in the EMR. Additionally, information on hospital discharge diagnoses coded in ICD-10 terminology is linked to patients included in BIFAP for a subset of periods and regions participating in the database. All information on prescriptions of medicines by the PCP is incorporated and linked by the PCP to a health problem (episode of care), and information on the dispensation of medicines at pharmacies is extracted from the e-prescription system that is widely implemented in Spain.

The project started in 2001 and the current version of the database with information until December 2020 includes clinical information of 14,810 primary care practices (PCPs) and paediatricians. Nine participant autonomous regions send their data to BIFAP every year. BIFAP database currently includes anonymized clinical and prescription/dispensing data from around 20 million (17 active population) patients representing 92% of all patients of those regions participating in the database, and 32% of the Spanish population. Mean duration of follow-up in the database is 9 years. From several regions, hospitalization data can be linked. In this study, this subpopulation is called BIFAP-HOSP-PC.

9.5.2 ES: SIDIAP (CATALUNYA)

The Information System for Research in Primary Care (Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP) in Catalonia, Spain, is a primary care database set up by the Institute of Research in Primary Care (Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina [IDIAP JGol]) and Catalan Institute of Health (Institut Català de la Salut). [ICS]). The database collects information from 278 primary health care centres and includes more than 5.8 million patients covered by the Catalan Institute of Health (approximately 78% of the Catalan population) and is highly representative of the Catalan population.

SIDIAP data comprise the clinical and referral events registered by primary care health professionals (i.e., GPs, paediatricians, and nurses) and administrative staff in electronic medical records, comprehensive demographic information, community pharmacy invoicing data, specialist referrals, and primary care laboratory test results. SIDIAP can also be linked to other data sources, such as the hospital discharge database, on a project-by-project basis. Health professionals gather this information using International Classification of Diseases, 10th Revision (ICD-10) codes, ATC codes for medicines, and structured forms designed for the collection of variables relevant to primary care clinical management, such as country of origin, sex, age, height, weight, body mass index, tobacco and alcohol use, blood pressure measurements, and blood/urine test results. In relation to vaccines, information on all routine childhood and adult immunisations is included in addition to the antigen and the number of administered doses.

SIDIAP was characterised in the IMI-ADVANCE project and considered fit for purpose for vaccine coverage, benefits, and risk assessment. An algorithm to identify pregnancies

has been previously used within SIDIAP. The algorithm uses diagnosis codes recorded in primary healthcare records during pregnancy and information recorded in the sexual and reproductive healthcare registries, including LMP, gestational week, expected date of delivery, actual date of delivery or termination, and pregnancy outcomes. Approximately 50% to 60% of pregnant women in Catalonia are attended in the sexual and reproductive healthcare centres that contribute data to SIDIAP. Approximately 70% of infant records can be linked to maternal records and used for research. The protocol will be evaluated by the SIDIAP Scientific Committe and by the IDIAPJGol Ethics Committee, the approval can take up to 4 weeks. The timeframe for data availability after the approval by the two local Committees is one month.

9.5.3 ES: VID (VALENCIA)

The Valencia health system integrated database (VID) is a set of multiple, public, population-wide electronic databases for the Valencia Region, the fourth most populated Spanish region, with ≈5 million inhabitants and an annual birth cohort of 48000 new-borns, representing 10.7% of the Spanish population and around 1% of the European population. The VID provides exhaustive longitudinal information including sociodemographic and administrative data (sex, age, nationality, etc.), clinical (diagnoses, procedures, diagnostic tests, imaging, etc.), pharmaceutical (prescription, dispensation) and healthcare utilization data from hospital care, emergency departments, specialized care (including mental and obstetrics care), primary care and other public health services. It also includes a set of associated population databases and registries of significant care areas such as cancer, rare diseases, vaccines, congenital anomalies, microbiology and others, and public health databases from the population screening programmers. All electronic health systems in the VID use the ICD-9-CM and the ICD-10-CM. All the information in the VID databases can be linked at the individual level through a single personal identification code. The databases were initiated at different moments in time, but all in all the VID provides comprehensive individual-level data fed by all the databases from 2008 to date. Information on PCR test results as well as serological/antibody tests results for the whole population of the Valencia region is available and linkable from the Microbiological Surveillance Network (RedMIVA). The Foundation for the Promotion of Health and Biomedical Research of Valencia Region (FISABIO) is Data Access Provider for Valencia Integrated Databases (VID).

9.5.4 ES: EPICHRON (ARAGON)

The EPICHRON database links sociodemographic and clinical anonymised information from 2010 to present for all the users of the public health system in Aragón (approximately 98% of the reference population). This database is built from the BIGAN platform, which integrates a technical infrastructure and a data lake gathering individual patient data from the regional health service information systems, including primary care, specialised care, hospitalisations, emergency department visits, drug prescriptions, image diagnosis, laboratory tests, diagnostics, vaccination, medical history, and demographics from the users of the public health system of Aragon, which comprises about 2 million individuals historic data and an active population of 1.3 million individuals.

9.5.5 IT: PEDIANET

Pedianet is a national population database that contains anonymous patient-level data of more than 500,000 children since 2004, corresponding to around 4% of the annual

paediatric population who received healthcare from family paediatricians (FPs) in Italy who were part of the PEDIANET network.
The network links FPs distributed throughout several Italian regions designated by the Italian NHS, including Friuli-Venezia Giulia, Liguria, Lombardia, Piemonte, Veneto, Lazio, Marche, Toscana, Abruzzo, Campania, Sardegna, and Sicilia, and who use the same software (Junior Bit®) (Padova, Italy) in their professional practice. Only children in Friuli Venezia Giulia can be linked to the immunization registry.
According to the Italian NHS, each child is assigned to a FP, who is the primary referral for health-related matters. In Italy, there is a tax-funded public healthcare system with universal access, and patients do not incur direct costs related to primary care visits. The Pedianet database captures several types of patient-level information, including the reason for accessing healthcare, health status, demographic data, diagnosis and clinical symptoms (free text or ICD-9-CM codes), drugs (ATC codes), specialist appointments, diagnostic procedures, hospital or emergency room (ER) admissions, growth parameters, and clinical outcome data. Informed consent is required from children's parents to enter the data in the database. The data collected from the child's parents/tutors by paediatricians enters the dedicated cloud already encrypted and anonymised. Pedianet researchers do not know the process to anonymise the data and cannot know the owner of the data in any way.

9.5.6 DK: DANISH NATIONAL REGISTRIES (DHR)

All Danish registries used in this study have a nationwide coverage and an almost 100% capture of contacts covering information on currently 5.9 million inhabitants plus historical information. Unambiguous person-level linkage across all data sources is possible via a unique identifier used in all Danish public records. Linked data from the following registries are available for the current project: the Danish Civil Registration System (identifier for linkage, age, sex, births, deaths, migrations); the Danish National Prescription Registry (outpatient dispensing in community pharmacies, no data on drugs administered in hospitals); the Danish National Health Service Register (GP contacts including vaccinations other than COVID-19); the Danish National Patient Registry (diagnoses and procedures from all hospital encounters); the Danish Vaccination Register (COVID-19 vaccinations only). Data are linked using a unique pseudonymized identifier on the servers of the Danish Health Data Authority (SDS). Individual-level data will be analysed by uploading and running of analytic scripts on the SDS servers and aggregate data that does not allow backtracking to individuals in accordance with the data regulation will be used for reporting. The Danish national registries are listed as a resource in the Catalogue of RWD sources and studies by EMA.

9.5.7 NO: NORWEGIAN NATIONAL LINKED REGISTERS AT UIO (NHR)

The core data that the University of Oslo (UiO) has access to are the health care administrative data banks of the entire Norwegian population, which amounts to approximately 5.3 million inhabitants. Norway has a universal public health care system, consisting of primary health care services and specialist healthcare services. Many population-based health registries were established in the 1960s, with use of unique personal identifiers facilitating linkage between registries. The mandatory national health registries were established to maintain national functions. They are used for health analysis, health statistics, improving the quality of healthcare, research, administration and emergency preparedness. The Norwegian data sources used in this project are the

national, mandatory Norwegian Surveillance System for Communicable Diseases (MSIS), which will be linked to five national health registries, i.e. the Medical Birth Registry, the National Patient Register, Norway Control and Payment of Health Reimbursement, the Norwegian Immunisation Registry, and the National Prescription Registry.
Information about all Norwegian National Registries can be found here: www.fhi.no/en/more/access-to-data/about-the-national-health-registries2/.
In this project, University of Oslo is Data Access Provider for Norwegian national registry data. Their current Norwegian health registry data will be used, capitalizing on the existing ETL's and quality checked data instance. In specific, UiO will contribute with ETL'd data on all residents in Norway between 1.1.2017- 31.12.2022, with historical data on these individuals back to 2010. Consequently, we will not be able to provide analysis as a near real-time analysis. Some ICD-10 codes are not at the 4-digit level.

9.5.8 UK: CLINICAL PRACTICE RESEARCH DATABASE (CPRD)

The Clinical Practice Research Database (CPRD) from the UK collates the computerised medical records of GPs in the UK who act as the gatekeepers of health care and maintain patients' life-long electronic health records. Accordingly, GPs are responsible for primary health care and specialist referrals, and they also store information about specialist referrals and hospitalisations. General practitioners act as the first point of contact for any non-emergency health-related issues, which may then be managed within primary care and/or referred to secondary care, as necessary. Secondary care teams also provide information to GPs about their patients, including key diagnoses. The data recorded in the CPRD include demographic information, prescription details, clinical events, preventive care, specialist referrals, hospital admissions, and major outcomes, including death. Most of the data is coded using Read or SNOMED codes. Data validation with original records (specialist letters) is also available. The population in the data bank is generalisable to the UK population based on age, sex, socioeconomic class, and national geographic coverage CPRD Aurum versions is used in this project. There are currently approximately 59 million individuals (acceptable for research purposes) -17 million of whom are active (ie, still alive and registered with the GP practice)- in over 2,000 primary care practices (https://cprd.com/Data). Data include demographics, all GP/health care professional consultations (e.g., phone calls, letters, e- mails, in surgery, at home), diagnoses and symptoms, laboratory test results, treatments (including all prescriptions), all data referrals to other care providers, hospital discharge summary (date and Read/SNOMED codes), hospital clinic summary, preventive treatment and immunisations, and death (date and cause). For a proportion of the CPRD panel practices (> 80%), the GPs have agreed to permit the CPRD to link at the patient level to HES data. Access to CPRD data will be provided by University Utrecht.

9.5.9 FR: *SYSTÈME NATIONAL DES DONNÉES DE SANTÉ* (SNDS)

The *Système National des Données de Santé* (SNDS) is the French nationwide healthcare database. It currently covers the overall French population (about 67 million persons) from birth (or immigration) to death (or emigration), even if a subject changes occupation or retires. Using a unique pseudonymized identifier, the SNDS merges all reimbursed outpatient claims from all French health care insurance schemes (SNIIRAM database), hospital-discharge summaries from French public and private hospitals (PMSI database),

and the national death registry. SNDS data are available since 2006 and contains information on:

- General characteristics: gender, year of birth, area of residence, deprivation index, etc;
- Death: month, year and cause;
- Long-term disease registration associated with an ICD-10 diagnostic code;
- Outpatient reimbursed healthcare expenditures with dates and codes (but not the medical indication nor result): visits, medical procedures, nursing acts, physiotherapy, lab tests, dispensed drugs and medical devices, etc. For each expenditure, associated costs, prescriber and caregiver information (specialty, private/public practice) and the corresponding dates are provided;
- Inpatient details: primary, related and associated ICD-10 diagnostic codes resulting from hospital discharge summaries with the date and duration of the hospital stay, the performed medical procedures (but no results), lab tests (but no results) and the related costs. Drugs included in the diagnosis related group cost are not captured. However, expensive drugs (i.e., the one charged in addition to the group cost) are.

Outpatient data (SNIIRAM) are uploaded to the SNDS throughout the year. It is admitted that a lag of around 6 months is required to catch 90% of the dispensing. Inpatient data (PMSI) are uploaded in one time, at the end of the following year. Hence, we consider that complete SNDS data of year Y are available in January of the year Y+2.
SNDS access is regulated: each study involving the human person with or without data extraction from the SNDS needs approval from the *Comité Ethique et Scientifique pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé* (CESREES) in charge of assessing scientific quality of the project, and authorization from the *Commission Nationale de l'Informatique et des Libertés* (CNIL) which is the French data protection authority, and then an agreement with the SNDS data holder (CNAM) for data extraction.

9.5.10 FI: FINNISH NATIONAL REGISTERS (FNR)

Finnish national data registers account for a total population of 5.4 million inhabitants. Main linkable data banks are: 1. *Hospital discharge register*: use of in- and outpatient services. Diagnoses for each admission are made by the attending physician. The register contains the following information on each hospital visit: dates, reason for hospital stay, specialty of the caring unit, date of operation, up to five operational codes (NOMESCO classification), where the patient was discharged to and assessment of need for assistance in activities of daily life. Since 2009, the data bank contains outpatient visits to specialised healthcare and since 2011 to primary healthcare. Laboratory and physiological measurements are available since 2015. 2. *Kanta electronic prescriptions:* all prescribed medicines purchased by an individual. Medicines used in hospitals are not included, but the register covers prescriptions written by hospital physicians and dispensed in community settings. Data on dispensing date, number of packages, tablets and defined daily dose (DDD) are available. Medicines are classified according to Anatomical Therapeutic Chemical (ATC)–classification system. 3. *Special reimbursement register:* entitlement to special reimbursement due to severe chronic diseases such as Alzheimer's disease, diabetes, psychosis, epilepsy, asthma, c*hronic obstructive pulmonary disease and* several cardiovascular diseases. The diagnoses are based on explicit predefined criteria.

4. *Statistics Finland* is the statistical authority of Finland, producing the majority of official statistics and conducting the population census, which has solely been based on the register data since 1990. These censuses include indicators of socioeconomic position (e.g. education, occupational status and taxable income). The causes of death register are compiled from death certificate data containing underlying, direct, intervening, and contributing causes. Death certificates are issued by physicians and if an autopsy is required, by a medicolegal officer.

Table 5Table 5

Table 5 Data provenance and vocabulary for each SAFETY-VAC data source

Data Access Provider	Data source	Data banks available for this study	Vocabularies	Data update frequency
AEMPS	BIFAP	Primary care record, hospital discharge diagnosis, community pharmacy dispensing, date of death.	ICPC2, ICD9 and SNOMED for diagnosis. ATC for medicines.	2 months.
IDIAP J Gol	SIDIAP	Primary care record, outpatient specialist record, outpatient laboratory results, surveillance data, emergency room, hospital discharge diagnosis, long term facility diagnosis, date of death.	ICD10-CM for diagnosis. ATC for medicines. ATC and antigen for vaccines. ICD10-PCS for procedures.	6 months
FISABIO	VID	Primary care record, outpatient specialist record, outpatient laboratory results, surveillance data, emergency room visits, hospital discharge diagnosis, in-hospital prescribing, pharmacy dispensing outpatient, in-hospital prescription/dispensing, long term facility diagnosis, date and reasons of death.	ICD10-CM and ICD9-CM for diagnosis and procedures. ATC for medicines. Disease + text information for vaccines.	Instantaneous for outpatient data, every 6 months for inpatient data.
IACS	EPICHRON	Primary care record, outpatient laboratory results, emergency room visits, hospital discharge diagnosis, pharmacy dispensing outpatient, date of death.	ICPC, ICD9-CM and ICD10-CM for diagnosis. ATC for medicines. ICD10-CM for procedures.	3-6 months
SOSETE	PEDIANET	Primary care record, outpatient specialist diagnosis, surveillance data, emergency room visits, hospital discharge diagnosis, in-hospital prescribing (free text), outpatient prescription, date of death, reasons of death.	ICD9-CM and free text for diagnosis. ATC and free text for medicines. ATC and free text for vaccines. ICD9-CM and free text for procedures.	6 months
Utrecht University	CPRD-Aurum	Primary care diagnoses, prescriptions, lab tests, hospital admissions and procedures CPRD death date	Read/Snomed for primary care diagnoses, BNF/product codes, but we have linked to ATC. ICD-10 for hospital diagnoses, OPCS for hospital procedures	New release scheme of primary care is quarterly. Lag time of hospital data (HES) difficult to say, currently available until 03/2021, used to annually updated.
Aarhus University	Danish registries	Outpatient specialist diagnosis, laboratory results (hospital-based), emergency room visits, hospital discharge diagnosis, outpatient pharmacy dispensing, in-hospital prescription/dispensing, date of death, reasons of death (2 years lag time).	ICD-10 Danish modification for diagnosis. ATC and hospital internal codes for medicines. Internal code for vaccines. NOMESCO for procedures.	Depends on data source.
University of Eastern Finland	Finnish registries	Primary care record (with some restrictions), outpatient specialist diagnosis, outpatient laboratory results, surveillance data, emergency room visits, hospital discharge diagnosis, in-hospital laboratory results, outpatient pharmacy dispensing, long term facility diagnoses, date and reasons of death.	ICD-10 for diagnosis. ATC for medicines. ATC and free text for vaccines. NOMESCO for procedures.	Depending on data source, from 1 month to 1 year.

Data Access Provider	Data source	Data banks available for this study	Vocabularies	Data update frequency
BPE	SNDS	Outpatient healthcare (no results, no indication), pharmacy dispensing (quantity, dosage, name, no indication), public/private hospital stays with discharge diagnosis (no results), public hospital visits (no results, no indication), emergency room visits (with diagnosis if > 1 day, without if <=1 day), in-hospital dispensing/prescription (only for out-of-DRG drugs), date of death, reason of death.	ICD-10 for diagnosis. ATC for medicines and vaccines. CCAM for procedures, NABM for lab tests, LPP for (para)medical devices	1 year.
University of Oslo	Norwegian registries	Primary care record, outpatient specialist diagnosis, surveillance data (infectious diseases), emergency room visits, hospital discharge diagnosis, outpatient pharmacy dispensing, date and reasons of death.	ICPC, ICD10CM ATC for medicines	Depends on data source and waiting time

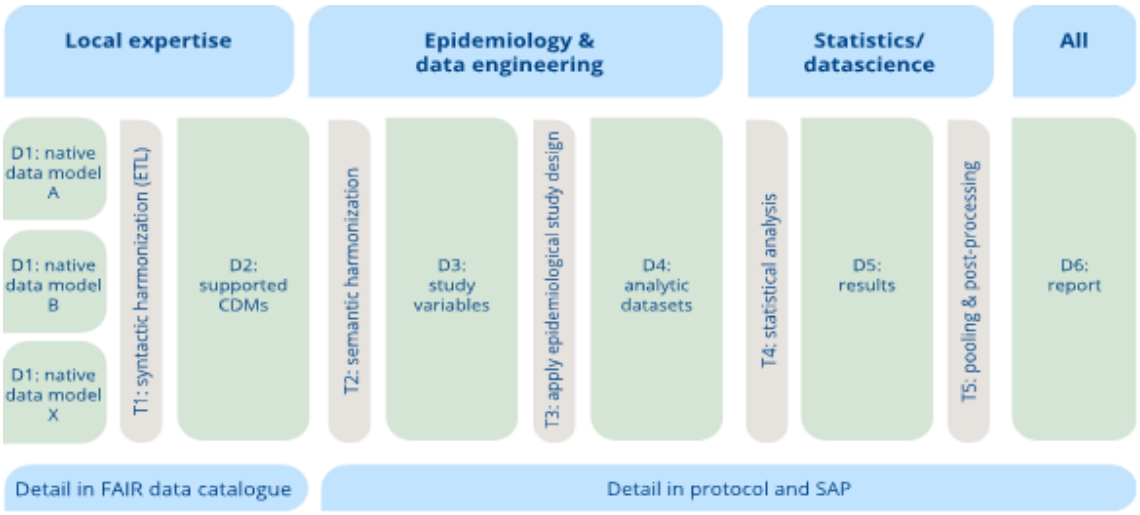
9.6 Study size

The study included all subjects eligible in the data sources that could be utilized for this report, approximating more than 40 million patients across 7 data sources.

9.7 Data transformation

The study was conducted in a distributed manner using the UMCU and VAC4EU tools, procedures, and pipeline. This pipeline can be viewed from a programming perspective (see Figure 2) or from a tool perspective (Figure 3). Figure 2 specifies the datasets (D) and transformation processes (T); programming follows this pipeline, with involvement of different types of experts.

Figure 2 Data management form the data transformation perspective.



D1: Original data can be in any native format

The RWD-RWE pipeline provided by VAC4EU and EU PE&PV starts with data banks that are controlled by the DAP, which can be in any format and are stored locally. The ETL template specification is shared in a searchable FAIR VAC4EU catalogue. The VAC4EU FAIR Molgenis data catalogue is a metadata management tool designed to contain searchable metadata describing organisations that can provide access to specific data sources.

T1: Syntactic harmonisation (ETL)

T1: Syntactic harmonisation is conducted through an extraction, transformation, and loading (ETL) process of native data into the requested CDM. To harmonise the structure of the data sets stored and maintained by each data partner, a shared syntactic foundation is used. The ETL process has various structured steps:

- DAPs are asked to share the data dictionaries of their data banks (selected tables and variable names/structure)
- Metadata (descriptive data about the data sources and data banks) & data dictionaries, are uploaded in the VAC4EU metadata catalogue
- DAPs make an ETL design

- The design is reviewed
- ETL is deployed

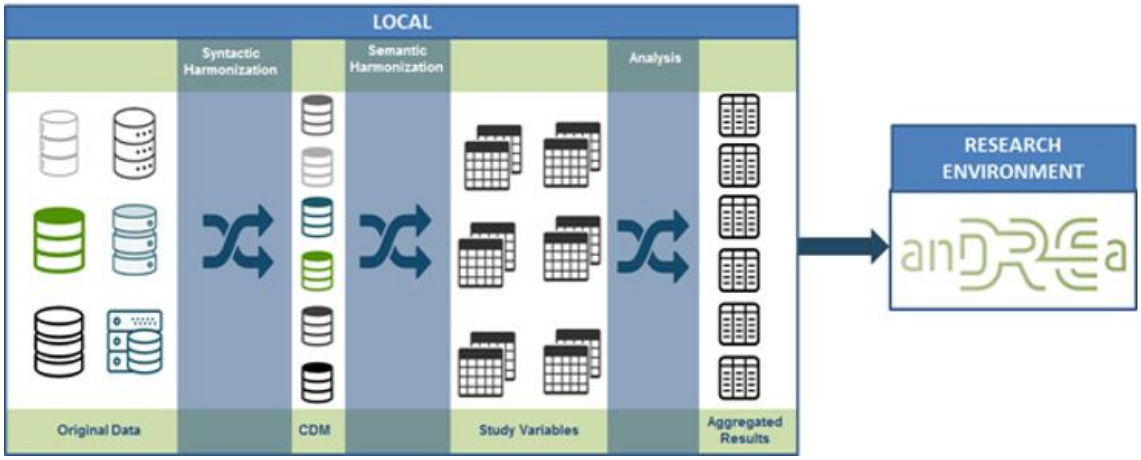
D2: Common data model

For this project we used ConcePTION CDM version v2.2. In the ConcePTION CDM the data is only syntactically harmonised, allowing for the data to remain in its original language (e.g., presence of different medical diagnostic systems such as ICD-9, ICD-10, SNOMED etc.).

T2: Semantic harmonisation

In this step we conduct time anchoring (observation periods, look back periods), clean the data such as the dose of vaccines, sort on record level, aggregate across multiple records, and combine concepts for implantation of algorithms, and rule-based creation of study variables. Based on the relevant diagnostic medical codes and keywords, as well as other relevant concepts (e.g., medications), one or more ***phenotype algorithms*** are constructed to operationalise the identification and measurement of each event or covariate.
In this phase of creation of study variables, semantic mapping is also implemented. This semantic mapping across different diagnostic vocabularies is conducted as part of the R study scripts. Machine readable code lists generated by the code list task force are ingested by the script. This is combined with the BRIDGE metadata file that defines time anchoring of the study variables (10).

Figure 3 Data management from a systems' and location's perspective.



D3: Study variables

D3 datasets are interim data sets with information on study variables for each study participant, the unit may be a person, a medicine or vaccine record, or episode of time. The design of these datasets is described in codebooks.

T3: application of epidemiological design

In the T3 step, epidemiological designs are applied such as sampling, matching (on specific variables and/or propensity scores) and selection based on inclusion and exclusion criteria using the study variables in the D3 datasets. The designs will be

implemented for the various study objectives using R scripts, and these may use the existing functions (R-cran) or functions that have been developed in the VAC4EU community.

D4: Analytical data set

D4 is an analytical dataset, and multiple D4 data sets may be produced based on the objectives of the study. The format is described initially in a codebook for communication between programmers and statisticians.

T4: Statistical analysis

This step in the data transformation pipeline produces statistical estimates such as descriptives (counts, percentages), distributions (mean, percentiles), rates (prevalence, incidence), regression coefficients, or other relevant estimates.

As per VAC4EU policy, the analytical code will publicly release with an open-source licence once the final report is made publicly available. The code will be published in this GitHub repository: https://github.com/VAC4EU/ROC18.

D5: Results

D5 is the set of estimands, tables or aggregate data that is transferred from the DAPs to the Digital Research Environment (DRE) (see Figure 3). The DRE is made available through UMCU. The DRE is a cloud-based, globally available research environment where data are stored and organised securely and where researchers can collaborate. All researchers who need access to the DRE received access to study-specific secure workspace by UMCU. Access to the workspaces is possible only after double authentication using an identification code and password together with the user's mobile phone for authentication. Downloading of files is possible only after requesting and receiving permission from a workspace member with an "privileged owner" role.

9.8 Statistical methods

The statistical analyses and methods utilized for this report are descriptive and comprise the INSIGHT data quality checks and indicators (described in section 10, with reference to dedicated Statistical Analysis Plans (SAPs)) plus the additional scripts to describe vaccine coverage, prevalence of covariates and incidence of outcomes. All analyses were conducted using R version R-4.03 or higher (Foundation for Statistical Computing, Vienna Austria).

9.8.1 VACCINE COVERAGE ESTIMATIONS

Vaccination coverage is the cumulative risk of being vaccinated with a particular antigen at a certain age. Vaccination schedules differ per country, but the WHO has benchmark data, related to certain antigens at certain ages. These data are used as an external benchmark and are provided in Table 5. To estimate coverage in a dynamic population is challenging because of left and right censoring. In the IMI-ADVANCE project methodological work was conducted to explore the best methods to estimate coverage in a dynamic population. We refer to Braeye et al. (11) for a simulation study that compared

different methods. The authors concluded that the inverse probability weighting (IPW)/Cumulative distribution function methods were generally the least biased. Preference for a specific method should be based on the type of censoring and type of dependence between completeness of follow-up and vaccination. A subsequent paper focused on childhood, adolescent and elderly and made recommendations on the methods to be used.

The following outcome parameters were estimated:

- Number of doses administered by vaccine during study period.
- Vaccine coverage curves (cumulative incidence) by birth year for childhood vaccines, for HPV from 9 years of age. Estimates have been provided at certain ages (see Table 6).

As explained, for the coverage estimation we followed the methodological approaches developed in the ADVANCE project (12). The Period Prevalence follow-up (PP-fu)-method relies on the assumption that the age-specific coverage estimated from the part of the population in follow-up at any age in weeks represent the age-specific coverage of the population. The IPW-method relies on the assumption that the proportion of persons in follow-up receiving a vaccine during a certain age equals the proportion of persons not in follow-up receiving a vaccine. Since this assumption is likely violated for the influenza-vaccine study population, as in older age groups death is a common cause of loss to follow-up, the IPW-method was not applied to influenza-vaccine. A general summary of these assumptions is that with the PP-fu-method we assumed that the observed coverage equalled the study population coverage, while with the IPW-method we estimate the coverage. The IPW-method accounts for both left and right censoring of vaccinations, but can produce unstable estimates when weights are very small or large and bias can accumulate as the method sums over the weekly estimated number of vaccinations (12). In this study we applied the IPW method for early childhood vaccinations, since we conditioned on start of birth for left censoring and addressed right censoring with IPW. For HPV vaccination and influenza vaccine, we applied the PP_{FU} method due the potentially high proportion of incomplete follow-ups over a longer period.
For childhood vaccines (see Table 1) coverage was estimated by birth year over age in months using only those persons that were born and in follow-up during the study period. The number of persons in follow-up for at least one day during an age in months was counted. Then, the number of persons who received a vaccination during that month was counted as well, and those who had a registered vaccination during that age-month. A letter (A, B, C, D, E) was assigned to every age-month of every person.

$$A_i = in\ follow{-}up\ (FU)\ during\ age\ i,\ vaccinated\ during\ age\ i$$

$$B_i = in\ FU\ during\ age\ i,\ vaccination\ recorded\ before\ age\ i$$

$$C_i = in\ FU\ during\ age\ i,\ no\ recorded\ vaccination\ before\ age\ i$$

$$D_i = Not\ in\ FU\ during\ age\ i,\ vaccination\ recorded\ before\ age\ i$$

$$E_i = Not\ in\ FU\ during\ age\ i,\ no\ recorded\ vaccination\ before\ age\ i$$

From the data aggregated by birth year, we calculated and produced coverage curves applying the following methods:

Period Prevalence: Follow-Up (PP_{FU})

The PP_{FU} estimate for month *i* is the number of vaccinated persons in follow-up divided by the number of persons in follow-up during month *i*.

$$PP_{FU,i} = \frac{A_i + B_i}{A_i + B_i + C_i}$$

$PP_{FU,i}$ calculates the vaccination coverage by dividing the number of vaccinated persons in follow-up prior and during month *i* over the total number of persons in follow-up during month *i*.

Inverse probability weighted (IPW)

When using the IPW method, to address right censoring inverse probability weighting is applied

$$IPW_i = \frac{\sum_{0 \rightarrow i} VaccIPW,_i}{NinFU}$$

IPW_i is the coverage estimated by the IPW_i-method during month *i*.

$$VaccIPW,_i = \frac{VACCobserved,i}{FU_{proportion,\,i}}$$

$VaccIPW,_i$ is the estimated number of vaccinations during month *i*; *VACC observed, i* is the observed number of vaccinations during month *i*; and $FU_{proportion,i}$ is the proportion of persons in follow-up during month *i*. $NinFU$ is the total number of persons in the birth cohort.

The age at which the coverage has been estimated, the method to estimate coverage, and the benchmark indicator per vaccine are provided in Table 6.

Table 6. Vaccine, age of dose assessment (months), method and main reference indicator for the selected vaccines

Vaccine	Dose 1 (months)	Dose 2 (months)	Dose 3 (months)	Method to estimate coverage	WHO/ECDC assessment indicator
Measles-Mumps-Rubella	12	23		IPW	MCV1 (13) (Measles-containing-vaccine first dose)
Diphtheria-Pertussis-Tetanus	12	12	23	IPW	DTP3 (14) (Diphteria tetanus toxiod and pertussis, dose 3)
Hib	12	12	23	IPW	Hib3 (15) (Haemophilus influenzae type B, dose 3)
Hepatitis B	12	12	24	IPW	HepB3 (16) (Hepatitis B, dose 3)
Polio	12	12	23	IPW	Pol 3 (17) (Polio, dose 3)
Pneumococcal	12	12		IPW	PCV (18) (Pneumococcal conjugate vaccines, dose 2)
Influenza	Yearly			PP_{FU}	n.a.
Varicella	24			IPW	n.a.
HPV	14 years, women	14 years, women		PP_{FU}	HPV (19) (HPV, dose 2, in women)
Rotavirus	12	12		IPW	RotaC (20) (Rotavirus, dose 2)
Tuberculosis	24			IPW	BCG (21) (Bacille Calmette-Guérin (BCG) vaccine, dose 1)
Meningococcal	15	15		IPW	n.a.
Coronavirus				PP_{FU}	Covid-19 (22) (we reported the coverage in age band 60+)

In order to assess completeness of vaccination data in the specific data sources, we compared the coverage estimates with the most recently published estimates from WHO. For COVID-19 vaccines, we used data from the ECDC. A priori, we decided that if coverage estimates in the databases deviate more than 10% (relative) from reference data, this is of concern. Of note, coverage reported to WHO may have varying origin, birth cohorts and years of assessment, as described in the links. We have used the most recent data available, but changes during lockdown were not considered.

9.8.2 COVARIATE PREVALENCE

Prevalence of covariates was measured at the start of follow-up and prior to start of COVID-19 vaccine roll-out. In this report, it was decided to report on prevalence of covariates for the latter since these can be benchmarked with data from the CVM study (22). Covariates were assessed within a lookback of 365 days prior to 01-12-2020 for diagnoses codes and medicines.

9.8.3 INCIDENCE & PREVALENCE RATES OF OUTCOMES

Incidence rates of events were calculated based on the first occurrence of an event and requiring absence of that event in the year prior (population at risk). Upon occurrence of the event, follow-up time was censored, see Figure 1.
Point prevalence estimates were calculated at the start of each year as follows: the numerator were the persons with the disease in the year prior, and the denominator were all persons present at the start of each calendar year.

One-year period prevalence estimates have been calculated as follows: the numerator comprised all persons who at the start of the year either had the disease in the year prior or developed the disease during the calendar year. The denominator comprise the person-years of follow-up in that year as an estimate of the average number of patients in that year (to avoid immortal time bias, we did not condition on being fully available). All estimates were age-standardised to the Eurostat population. Age-specific estimates were calculated in the following categories: 0-1, 2-4, 5-11, 12-17, 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, and 80+ years.

9.9 Quality control

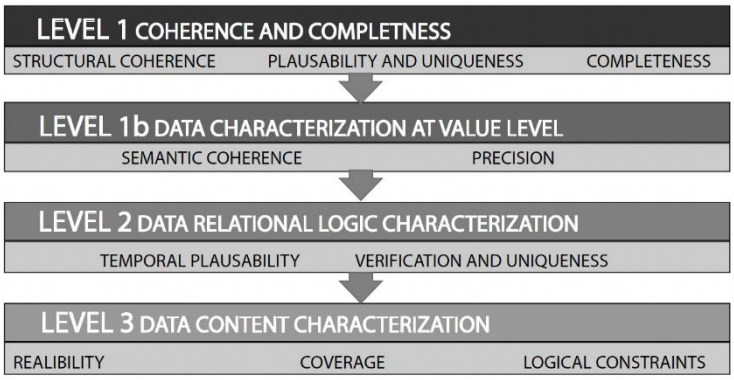
The INSIGHT data quality assessment R-tool for data converted into the ConcePTION CDM allows for a detailed characterization of the data source instance that is used for this study, including an overview of the anticipated availability and quality of exposure to selected vaccines of interest. INSIGHT is a public set of R scripts that identifies potential data quality issues in ConcePTION CDM-standardized instances through the systematic execution and summary of over 588 configurable data quality assessments (23). All INSIGHT scripts are publicly available on https://github.com/UMC-Utrecht-RWE.

For the INSIGHT level 1-3 quality checks, detailed statistical analysis plans are available on public repositories:

- https://github.com/UMC-Utrecht-RWE/INSIGHT-Level1 (24)
- https://github.com/UMC-Utrecht-RWE/INSIGHT-Level2 (25)
- https://github.com/UMC-Utrecht-RWE/INSIGHT-Level3 (26)

Level 1 focuses on compliance with the ConcePTION CDM specifications and data completeness. Level 2 evaluates the temporal plausibility of events and the uniqueness of records. Level 3 provides an overview of distributions, outliers, and trends over time. The data quality assessments are run locally by the DAP and assessed centrally by a data quality revisor together with the DAP's representatives.
INSIGHT is a tool that aligns with and operationalizes the five dimensions of the EMA data quality framework: reliability, extensiveness, coherence, timeliness, and relevance (Figure 4). Data quality is the sum of several internal and external features of data. An important feature of the VAC4EU procedures is that each data instance (a version of the original data that is converted into the CDM) will undergo through the quality assessment using the INSIGHT pipeline, prior to running an analysis study script, to ensure quality of data, since this may vary largely between different instances.

Figure 4. Hierarchy and dimensions of data quality assessment in the INSIGHT tools mapped to the EMA data quality framework.



To ensure that data quality indicators can be inspected, results are presented in a HTML format for each level, facilitating their understanding and sharing. These reports contain summary tables that allow for a concise representation of data quality indicators and graphs that provide a visual representation of trends and patterns. The INSIGHT data quality assessments are an iterative process for each data instance. Each level can be rerun until the required quality is attained or all constraints are noted.

After the quality indicators of the data instances and outputs of the analytical scripts have been generated, we performed the fitness-for-purpose of the data instance using, implementing, and adapting "The Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework" (SPIFD) from Gatto et al (27). This is an assessment tool aimed at conducting feasibility assessment to determine whether a data source is fit-for-purpose for specific real-world effectiveness and safety studies. The SPIFD framework is composed of three operative steps: i) operationalization and ranking of minimal criteria required to answer the research question; ii) identification and narrowing down data sources options, and iii) conducting detailed feasibility assessment. The last allowed us to tabulate different evaluation items and therefore to score them. Under this approach we could produce a list of data sources that could be included in the network of real-world data sources for optional vaccine safety studies.

10 RESULTS

10.1 Data quality, vaccine coverage and outcome rates (Objective 1a).

Table 7 presents an overview of the data instances used and quality levels that were checked. For this report dated 27 May 2024, BIFAP, SIDIAP, VID, Pedianet, DHR, NHR, CPRD data were available. Results on incidence and coverage could not be generated yet for Epichron (the permission to produce results has not been granted yet for current study), Finland (delay of data extraction) and France (no permission to produce results for current study), but for Epichron and SNDS quality checks on available data instances were available.

Table 7. Information on the data instance, the conducted level checks for that instance and the additional analysis on incidence and coverage for the feasibility assessment per data source.

N	Country	Data Source	Data Instance (recommended end date)	INSIGHT Quality Control Framework				Analysis script of incidence, prevalence and coverage
				Level 1	Level 1b	Level 2	Level 3	
1	Spain	BIFAP	30 Apr 2022	Yes	Yes	No	No	Yes
2	Spain	SIDIAP	30 Jun 2023	Yes	Yes	Yes	Yes*	Yes
3	Spain	VID/FISABIO	22 Mar 2022	Yes	Yes	Yes	Yes	Yes
4	Italy	PEDIANET	31 Dec 2022	Yes	Yes	Yes	Yes	Yes
5	Denmark	DHR	31 Dec 2022	Yes	Yes	Yes	No	Yes
6	Norway	NHR at UIO	14 Jan 2023	Yes	Yes	Yes	No	Yes
7	UK	CPRD	21 Mar 2022	Yes	Yes	Yes*	No	Yes
8	Spain	Epichron	31 Jul 2023	Yes	Yes	Yes	No	No
9	Finland	FR	31 Dec 2023**	n/a	n/a	n/a	n/a	No
10	France	SNDS	31 Dec 2020	Yes	Yes	Yes	Yes	No

*Pending for final approval**Data requested n/a: not applicable

10.1.1 QUALITY CHECKS

10.1.1.1 Level 1 and level 1b: Coherence and completeness

Of the 10 data sources, 9 ran and passed Level 1 and 1b checks. Full HTML Outputs are available on the Digital Research Environment. The quality checks outputs confirming the presence of the tables and variables needed to fill the corresponding CDM tables can be found in Tables A1 and A2 from **Annex 1**.

In Table 8, we present a summary of some quality indicators from INSIGHT Level 1 checks. Further information on Level 1 quality checks and graphical outputs can be found in Figures A1, A2 and A3 in **Annex 1**.

Finally, INSIGHT Level 1 also provides an overview of data availability of the source data, at table and variable level. An example can be found in Table A3 in **Annex 1**.

Table 8. Summary of some L1Level 1 checks data quality indicators.

		ES-BIFAP	ES-SIDIAP	ES-VID	ES-Epichron	IT-PEDIANET	FR-SNDS	DK-DHR*	NO-NHR	UK-CPRD
Completeness of information (%)	Birth date elements recorded and rounding of dates/months	day, month and year	day, month and year	day, month and year	day, month and year	day (15), month and year	year	day, month and year	month and year	day, month and year
	Person with a recorded death date	8%	8%	5%	2%	0%	2%	2%	3%	8%
	Sex recorded	100%	100%	100%	100%	100%	100%	100%	100%	100%
	Recorded country of birth	0%	24%	2%	0%	0%	0%	0%	100%	0%
	Recorded events codes	100%	100%	100%	100%	100%	100%	100%	100%	100%
	Recorded codes for medicines records	ATC code (100%), MPID (100%)	ATC code (100%), MPID (100%)	ATC code (100%), MPID (0%)	ATC code (100%), MPID (100%)	ATC code (100%), MPID (100%)	ATC code (98%), MPID (98%)	ATC code (100%), MPID (0%)	ATC code (100%), MPID (40%)	ATC code (100%), MPID (100%)
	Recorded odes to identify vaccine	ATC code (100%), Lot number (100%)	ATC code (63%), Vaccine type (100%)	Lot number (100%), Vaccine type (100%)	ATC code (99%), Vaccine type (100%), Lot number (35%)	ATC code (100%), Lot number (96%)	ATC code (100%)	Vaccine type*	ATC code (100%), Vaccine type (100%)	ATC code (100%)
	Vaccine dose recorded	80%	100%	100%	50%	8%	0%	100%	0%	100%
Provenance of event records		Primary care, Hospital	Primary care, Emergency, Hospital, Specialist and ICU	Primary care, Emergency, Hospital, Specialist and ICU	Primary care, Emergency, Hospital	Primary care, Emergency, Hospital	Primary care, Hospital	Emergency, Hospital, including outpatient specialist visits.	Primary care, Emergency, Hospital	Primary care
Provence of medicines records		90% dispensing, 15% prescription	100% Dispensing	100% Dispensing	100% Dispensing	100% Prescription	100% Dispensing	100% Dispensing	100% Dispensing	100% Prescription

ICU: intensive care unit; MPID: medicinal product identifier

*For DK-DHR the quality check outputs must be stored on the servers of the Danish Health Data Authority (SDS) and cannot be reproduced in a report, in compliance with the local regulations on data protection

10.1.1.2 Level 2: Logical checks

Nine data sources ran and passed Level 2 checks. The purpose of INSIGHT level 2 is to confirm temporal plausibility of data and assess uniqueness (23). Its outputs provide an overview of the logical relationship and integrity of values within and between variables and tables. Such metrics are in line with the dimensions of coherence and some complex metrics from the plausibility dimension as reported by the EMA data quality framework (28). INSIGHT Level 2 quality checks are divided in 8 sections, 4 of which are mandatory. The mandatory sections include the detection of date values before birth (Level 2.1), the detection of date values after death (Level 2.2), the detection of date values outside the observation period (Level 2.3) and a check on consistency of PERSON ID presence between the PERSONS CDM tables and other tables in the CDM (Level 2.4). A percentage of less than 5% discordance is conveyed to be acceptable.

SIDIAP, BIFAP, VID, CPRD and NHR had ≤1% of records before date of birth. In contrast, Epichron had up to 1.7% of dates relating to pregnancy-related diagnoses before birth. PEDIANET reached up to 4.8% of start dates of diagnoses before birth, because of rounding of birthdates.

None of the DAPs had records after death. Date values outside the observation periods were more frequent. While SIDIAP had no record outside the observation period, VID and CPRD reached a 14% for emergency room procedure dates and 8.4% for diagnoses in primary care and medicine dates of prescription, respectively. Epichron, PEDIANET and NDR had a few values outside observation period but always below 4%, and BIFAP 8.7% for records in survey tables. As for "PERSON IDs not present in PERSONS" CDM table, ES-VID reached a 6.9% but only for events coming from the specialist diagnosis consults. Every other issue for any DAP remained below 4%.

For DK-DHR, the quality check outputs must be stored on the servers of the SDS and cannot be reproduced in a report, in compliance with the local regulations on data protection. They passed Level 2 quality checks.

In Table 9 we summarize the results of INSIGHT Level 2 checks per data source.

Table 9 Percentage ranges of selected issues found in INSIGHT Level 2 checks per sub-level and Data Access Partner (DAP).

Percentage ranges	ES-BIFAP	ES-SIDIAP	ES-VID	ES-EPICHRON	IT-PEDIANET	FR-SNDS	DK-DHR*	NO-NHR	UK-CPRD
2.1 Date values before birth	0-0%	0-0%	0-0.1%	0-1.8%	0-4.8%	0-0%	0.0%	0-1.1%	0-0.02%
2.2 Date values after death	0-0%	0-0%	0-0%	0-0%	0-0%	0-0%	0%	0-0%	0-0%
2.3 Date values outside observation periods	0-8.7%	0-0.1%	0-13.9%	0-2.3%	0-3.7%	0-11.7%	<5%	0-2.1%	0-8.4%
2.4 Person IDs not in PERSONS table	0-0%	0-0%	0-6.9%	0-2.1%	0-0%	0-0%	0%	0-0%	0-1.8%
CDM table with highest percentage of issues	SURVEY_OBSERVATIONS	MEDICINES	PROCEDURES	MEDICAL_OBSERVATIONS	EVENTS	EVENTS, SURVEY_OBSERVATIONS	MEDICAL_OBSERVATIONS	VACCCINES	MEDICINES

*For DK-DHR the quality check outputs must be stored on the servers of the Danish Health Data Authority (SDS) and cannot be reproduced in a report, in compliance with the local regulations on data protection.

10.1.1.3 Level 3: Data content characterization

Level 3 data quality indicators provide distributions of population, diagnoses, medicines, vaccines, lifestyle factors, pregnancy, and temporal trends over calendar time for each specific variable. The primary objective is to allow for inspection of temporal changes in population, follow-up, medicines, vaccines, and disease rates (23).
Three data sources completed Level 3 quality checks and 2 passed them. SIDIAP had pending clarifications about some patterns and time windows. The purpose of Level 3 data quality indicators is to provide counts, rates and time distributions over calendar time of population, diagnoses, medicines, vaccines, and lifestyle factors, among others. The primary objective is to allow for inspection of temporal changes, which can be compared between instances and between DAPs, but also against external benchmarks to verify their fitness-for-purpose for a specific research question.
Table 10 provides an overview of some selected quality indicators from Level 3 per data source. The exposure patterns look consistent across data instances and are clearly impacted by the COVID-19 lockdown period and vaccines commercialization. Lifestyle variables were scarce in the quality checked instance for SIDIAP, VID and PEDIANET, but all of them were able to capture vaccine information even at different levels or using different manners.
Further details on INSIGHT Level 3 quality checks together with graphical outputs can be found in Figures A4, A5, A6 and A7 in **Annex 1**.

Table 10 Comparison of selected Level 3 INSIGHT quality indicators.

	ES-SIDIAP	**ES-VID**	**IT-PEDIANET**	**FR-SNDS**
Population tree as expected for the data source	Yes	Yes	Yes	Yes
Temporal patterns in Event rates	Drop in rates in 2020 which returns to previous levels afterward*	Drop in rates in 2020 which returns to previous levels afterward*	Increase in rates from 2020 onwards	n.a.
Level of detail of ATC codes	ATC7 (100%)	ATC7 (100%)	ATC4 (0.1%), ATC5 (3.1%), ATC7 (96.8%)	ATC1 (<0.1%), ATC3 (<0.1%), ATC4 (0.4%), ATC5 (2.0%), ATC6 (0.3%), ATC7 (97.3%)
Temporal patterns in medicines rates	Drop in rates in 2020 which returns to previous levels afterward*	Drop in rates in 2020 which returns to previous levels afterward*	Increase from 2020 onwards	Drop in rates in 2020 which returns to previous levels afterward*
Vaccination exposure identification	ATC5 (17.2%), ATC7 (82.8%)	Vaccine type (100%, covid19 and influenza)	ATC5 (6.6%), ATC7 (93.4%)	ATC7 (100%)
Temporal patterns in vaccination rates	Peak after 2020	Peak after 2020	Peak after 2020	Drop after 2017

*Probably due to COVID-19 pandemic. n.a: not available

10.1.2 ATTRITION AND DEMOGRAPHICS OF STUDY POPULATION

The period of data availability for the data instance differed across data sources. Table 4 lists these end dates. For VID, BIFAP and CPRD data instances, the end date was between March and April 2022. For PEDIANET and DHR, it was December 2022. For NHR, the end data was January 2023, and for SIDIAP June 2023. The lag time for the data instances varied between 8-22 months.

The study population of subjects registered in the data sources after 1/1/2017 with complete data on date of birth and sex and at least one year of valid data which was required for this study comprised 53,283,613 subjects.

The application of the exclusion criteria led to a final study population that represents between 64% (BIFAP_PC_HOSP) to 97% (NHR) of the initial data sources' population (Table 11). The most impacting exclusion criteria was the exit from the data source before 1/1/2017 for most of the data sources, except for BIFAP_PC_HOSP subpopulation, where the main exclusion was no observation periods available, sex or birth date information or no dates of entry or exit in the instance. This is a particular case since observation periods in BIFAP_PC_HOSP refers to the presence of both PC and HOSP data.

The demographic characteristics of the study population are presented in Table 12. CPRD provides the largest study population (15,791,053) across the participating data sources whereas PEDIANET (restricted to paediatric population) is the smallest one (44,922). Sex is balanced across all data sources, with female participants ranging from 48-52%. Median age is also similar across data sources, ranging from 34 to 43 years old, except for PEDIANET.

Table 11. Attrition per data source.

Exclusion criteria	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
Persons in the data instance of the data source	16244090	16244090	5371422	7441114	46392	6585104	5824950	17666697
No observation period available, sex or birth date missing or absurd or no dates of entry or exit	62	5007210*	26403	0	0		3286	855
Exit from the data source before 1/1/2017	2083290**	387421**	92067**	466727	0	39566	5623	0
Less than 365 days history at any point in time after 1.1.2017	631940	445867	157871	198295	1470	205734	108178	1874789
Final study population	13528798	10403592	5095081	6776092	44922	6339804	5707863	15791053

*Available observation periods in ES-BIFAP-PC-HOSP refers to the presence of both PC and HOSP data, non-linked population is excluded in this step
** only data from 2018 onwards

Table 12 Demographic characteristics of the study population

	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
Study population	13528798 (100%)	10403592 (100%)	5095081 (100%)	6776092 (100%)	44922 (100%)	6339804 (100%)	5707863 (100%)	15791053 (100%)
Female	6982546 (52%)	5404850 (52%)	2591561 (51%)	3417558 (50%)	21755 (48%)	3174088 (50%)	2824946 (49%)	7909588 (50%)
Male	6546252 (48%)	4998742 (48%)	2503520 (49%)	3358534 (50%)	23167 (52%)	3165716 (50%)	2882665 (51%)	7881465 (50%)
Other							252 (<0.1%)	
Age (IQR) 25%-50%-75%	23-41-58	24-41-58	23-43-59	21-39-56	0.00-2.00-5.00	18-38-58	17-36-55	19-34-53
Age in categories								
0-1	632782 (4.7%)	459591 (4.4%)	190069 (3.7%)	379484 (5.6%)	22280 (50%)	493401 (7.8%)	435384 (7.6%)	988335 (6.3%)
2-4	374836 (2.8%)	280321 (2.7%)	138383 (2.7%)	191092 (2.8%)	8497 (19%)	185500 (2.9%)	193606 (3.4%)	584850 (3.7%)
5-11	957149 (7.1%)	719830 (6.9%)	369368 (7.2%)	483946 (7.1%)	13673 (30%)	477918 (7.5%)	465014 (8.1%)	1275312 (8.1%)
12-17	741570 (5.5%)	559513 (5.4%)	314152 (6.2%)	385788 (5.7%)	472 (1.1%)	417101 (6.6%)	391162 (6.9%)	904508 (5.7%)
18-29	1705971 (13%)	1307033 (13%)	593042 (12%)	916027 (14%)		996083 (16%)	901603 (16%)	2946922 (19%)
30-39	2037793 (15%)	1577293 (15%)	665068 (13%)	1061466 (16%)		724941 (11%)	752119 (13%)	2409860 (15%)
40-49	2176611 (16%)	1686969 (16%)	839165 (16%)	1076033 (16%)		806419 (13%)	764737 (13%)	2029847 (13%)

	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
50-59	1787962 (13%)	1390485 (13%)	732676 (14%)	828090 (12%)		794541 (13%)	677072 (12%)	1848915 (12%)
60-69	1317525 (9.7%)	1028618 (9.9%)	562336 (11%)	631924 (9.3%)		677442 (11%)	566622 (9.9%)	1358681 (8.6%)
70-79	953138 (7.0%)	736756 (7.1%)	414939 (8.1%)	445952 (6.6%)		515908 (8.1%)	369932 (6.5%)	927876 (5.9%)
80+	843461 (6.2%)	657183 (6.3%)	275883 (5.4%)	376290 (5.6%)		250550 (4.0%)	190612 (3.3%)	515947 (3.3%)

The prevalence of comorbidities for the population at 1/12/2020 is presented in Table 13**.** Cardio/cerebrovascular disease is most commonly observed across databases (from 17.5 to 24.5%), except in PEDIANET. It is followed by lipid abnormalities going from 11% in SIDIAP to 16% in VID, this covariate was not captured in the CPRD instance. Mental health diseases were 8 to 12%, except in CPRD (3%, as psychiatric conditions were not included in the CPRD data instance) and PEDIANET (0.1%). The prevalence of diabetes type 1&2 was around 5% across databases, except in PEDIANET, which is children only. Cancer prevalence ranged from 0.1% in PEDIANET to 3.4% in NHR. The prevalence of hypertension was low, since this was based on diagnosis codes only. The pregnancy algorithm was only run by VID, SIDIAP and NHR, and therefore is not available for several data sources. Several conditions have 0% prevalence, indicating that these diagnostic codes were not extracted for the data instance in data sources that require extraction from controller (ES-VID, DK-DHR, UK-CPRD).

Table 13. Prevalence of comorbidities at 1-12-2020 (lookback period 365 days)

	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
Study population	11827601 (100%)	9038207 (100%)	4880104 (100%)	5864586 (100%)	36587 (100%)	5746218 (100%)	5328274 (100%)	12237133 (100%)
Female	6139288 (52%)	4728817 (52%)	2485964 (51%)	2968662 (51%)	17710 (48%)	2891537 (50%)	2641540 (50%)	6070795 (50%)
Pregnancy	0 (0.0%)	0 (0.0%)	24192 (1.0%)	26902 (0.9%)	0 (0.0%)	49561 (1.7%)	47855 (1.8%)	0 (0.0%)
Pernicious anemia	229 (0.0%)	1200 (0.0%)	8516 (0.2%)	2383 (0.0%)	6 (0.0%)	0 (0.0%)	214 (0.0%)	0 (0.0%)
Sickle cell disease	1022 (0.0%)	971 (0.0%)	491 (0.0%)	261 (0.0%)	19 (0.1%)	0 (0.0%)	136 (0.0%)	2688 (0.0%)
Arrhythmia	53900 (0.5%)	93969 (1.0%)	157189 (3.2%)	80346 (1.4%)	47 (0.1%)	76349 (1.3%)	109477 (2.1%)	102618 (0.8%)
Coronary artery disease	35986 (0.3%)	66011 (0.7%)	110465 (2.3%)	37702 (0.6%)	0 (0.0%)	43545 (0.8%)	75913 (1.4%)	69037 (0.6%)
Cardiocerebrovascular disease	2892167 (24.5%)	2189025 (24.2%)	1256445 (25.7%)	1260797 (21.5%)	131 (0.4%)	1439400 (25.0%)	1185690 (22.3%)	2147121 (17.5%)
Heart failure	27983 (0.2%)	52717 (0.6%)	63166 (1.3%)	35086 (0.6%)	10 (0.0%)	28898 (0.5%)	32854 (0.6%)	70769 (0.6%)
Myocardial infarction	6454 (0.1%)	20315 (0.2%)	39095 (0.8%)	13898 (0.2%)	0 (0.0%)	9059 (0.2%)	16092 (0.3%)	19476 (0.2%)
Valvular heart disease	4457 (0.0%)	19588 (0.2%)	18832 (0.4%)	17749 (0.3%)	0 (0.0%)	0 (0.0%)	40607 (0.8%)	20291 (0.2%)
Celiac disease	1351 (0.0%)	1664 (0.0%)	4642 (0.1%)	2115 (0.0%)	52 (0.1%)	0 (0.0%)	5655 (0.1%)	0 (0.0%)
Crohn disease	885 (0.0%)	2582 (0.0%)	8488 (0.2%)	1670 (0.0%)	6 (0.0%)	11242 (0.2%)	11003 (0.2%)	0 (0.0%)

	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
Gallstones	14176 (0.1%)	25667 (0.3%)	31316 (0.6%)	23317 (0.4%)	0 (0.0%)	0 (0.0%)	8468 (0.2%)	26232 (0.2%)
Autoimmune hepatitis	184 (0.0%)	505 (0.0%)	1127 (0.0%)	338 (0.0%)	0 (0.0%)	0 (0.0%)	1439 (0.0%)	1025 (0.0%)
Inflammatory bowel disease	2136 (0.0%)	5034 (0.1%)	17535 (0.4%)	3818 (0.1%)	8 (0.0%)	25063 (0.4%)	24856 (0.5%)	0 (0.0%)
Chronic liver disease	1566 (0.0%)	20855 (0.2%)	35078 (0.7%)	28714 (0.5%)	0 (0.0%)	0 (0.0%)	12621 (0.2%)	23085 (0.2%)
Non-alcoholic fatty liver	<5 (~%)	10406 (0.1%)	18391 (0.4%)	18644 (0.3%)	0 (0.0%)	0 (0.0%)	1644 (0.0%)	14200 (0.1%)
Pancreatitis, chronic	729 (0.0%)	3108 (0.0%)	3803 (0.1%)	2765 (0.0%)	<5 (~%)	0 (0.0%)	3401 (0.1%)	816 (0.0%)
Ulcerative colitis	1055 (0.0%)	2299 (0.0%)	9821 (0.2%)	2187 (0.0%)	<5 (~%)	14160 (0.2%)	14802 (0.3%)	0 (0.0%)
Diabetes type 1&2	730640 (6.2%)	544642 (6.0%)	375156 (7.7%)	355271 (6.1%)	41 (0.1%)	274807 (4.8%)	229739 (4.3%)	606205 (5.0%)
Gout	11719 (0.1%)	12956 (0.1%)	47483 (1.0%)	8862 (0.2%)	0 (0.0%)	5635 (0.1%)	5075 (0.1%)	46558 (0.4%)
Lipid abnormalities	1664190 (14.1%)	1268842 (14.0%)	793130 (16.3%)	650700 (11.1%)	<5 (~%)	694684 (12.1%)	603981 (11.3%)	0 (0.0%)
Chronic renal disease	22095 (0.2%)	47222 (0.5%)	85826 (1.8%)	45556 (0.8%)	<5 (~%)	0 (0.0%)	29655 (0.6%)	18728 (0.2%)
Mild/moderate chronic renal disease	3417 (0.0%)	33238 (0.4%)	70432 (1.4%)	42677 (0.7%)	<5 (~%)	20131 (0.4%)	27559 (0.5%)	9176 (0.1%)
Hepatitis B	845 (0.0%)	2043 (0.0%)	2251 (0.0%)	2300 (0.0%)	0 (0.0%)	0 (0.0%)	3777 (0.1%)	1013 (0.0%)
HIV/AIDS	217 (0.0%)	1712 (0.0%)	15273 (0.3%)	753 (0.0%)	0 (0.0%)	4891 (0.1%)	7525 (0.1%)	2744 (0.0%)
Herpes simplex infection	19675 (0.2%)	15223 (0.2%)	21186 (0.4%)	12161 (0.2%)	74 (0.2%)	0 (0.0%)	1794 (0.0%)	35 (0.0%)
Influenza	167623 (1.4%)	112296 (1.2%)	41379 (0.8%)	71937 (1.2%)	3415 (9.3%)	n/a	588 (0.0%)	85830 (0.7%)
Infection	1748004 (14.8%)	1258629 (13.9%)	1267878 (26.0%)	1007519 (17.2%)	8060 (22.0%)	1283727 (22.3%)	938586 (17.6%)	2253634 (18.4%)
Respiratory infection	169507 (1.4%)	196841 (2.2%)	232317 (4.8%)	289268 (4.9%)	5016 (13.7%)	0 (0.0%)	11486 (0.2%)	85830 (0.7%)
Sepsis	585 (0.0%)	12527 (0.1%)	5884 (0.1%)	8270 (0.1%)	0 (0.0%)	0 (0.0%)	2715 (0.1%)	0 (0.0%)
Atopic dermatitis	141203 (1.2%)	103868 (1.1%)	1671 (0.0%)	90628 (1.5%)	664 (1.8%)	0 (0.0%)	37555 (0.7%)	0 (0.0%)
Dermatomyositis	0 (0.0%)	123 (0.0%)	456 (0.0%)	160 (0.0%)	0 (0.0%)	0 (0.0%)	497 (0.0%)	0 (0.0%)
Immunocompromised status	119189 (1.0%)	129053 (1.4%)	262403 (5.4%)	111586 (1.9%)	155 (0.4%)	179901 (3.1%)	204524 (3.8%)	164938 (1.3%)
Sjogren syndrome	343 (0.0%)	1134 (0.0%)	2398 (0.0%)	2146 (0.0%)	0 (0.0%)	0 (0.0%)	10254 (0.2%)	1621 (0.0%)
Systemic lupus erythematosus	640 (0.0%)	1626 (0.0%)	4738 (0.1%)	1417 (0.0%)	0 (0.0%)	2359 (0.0%)	2981 (0.1%)	2079 (0.0%)
Transplant recipient	12655 (0.1%)	9622 (0.1%)	6669 (0.1%)	9211 (0.2%)	<5 (~%)	212 (0.0%)	6178 (0.1%)	0 (0.0%)
Obesity	37106 (0.3%)	58227 (0.6%)	71861 (1.5%)	62161 (1.1%)	321 (0.9%)	41300 (0.7%)	47721 (0.9%)	16679 (0.1%)
Rheumatoid Arthritis	2089 (0.0%)	5732 (0.1%)	20327 (0.4%)	3723 (0.1%)	<5 (~%)	23366 (0.4%)	20419 (0.4%)	23756 (0.2%)
Alcohol abuse	3834 (0.0%)	22828 (0.3%)	33947 (0.7%)	21372 (0.4%)	0 (0.0%)	17849 (0.3%)	20477 (0.4%)	4515 (0.0%)
Sleep disorders	68717 (0.6%)	53512 (0.6%)	216413 (4.4%)	6443 (0.1%)	115 (0.3%)	0 (0.0%)	992 (0.0%)	312 (0.0%)
Mental health diseases	1218618 (10.3%)	923298 (10.2%)	576496 (11.8%)	603555 (10.3%)	50 (0.1%)	494278 (8.6%)	482475 (9.1%)	372214 (3.0%)
Dementia	6662 (0.1%)	26045 (0.3%)	67467 (1.4%)	28028 (0.5%)	0 (0.0%)	16101 (0.3%)	9594 (0.2%)	53516 (0.4%)
Myasthenia gravis	271 (0.0%)	674 (0.0%)	1639 (0.0%)	534 (0.0%)	0 (0.0%)	919 (0.0%)	770 (0.0%)	0 (0.0%)
Paralysis	360 (0.0%)	7299 (0.1%)	7783 (0.2%)	7029 (0.1%)	8 (0.0%)	0 (0.0%)	2012 (0.0%)	9 (0.0%)

	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
Malignancies	120408 (1.0%)	123131 (1.4%)	126758 (2.6%)	78285 (1.3%)	31 (0.1%)	106037 (1.8%)	178935 (3.4%)	129029 (1.1%)
Preeclampsia	597 (0.0%)	2234 (0.0%)	1593 (0.0%)	1967 (0.0%)	0 (0.0%)	3938 (0.1%)	3696 (0.1%)	1981 (0.0%)
Psoriasis	1947 (0.0%)	5279 (0.1%)	31964 (0.7%)	8868 (0.2%)	98 (0.3%)	8402 (0.1%)	31621 (0.6%)	0 (0.0%)
Hypertension	17089 (0.1%)	171877 (1.9%)	848378 (17.4%)	162487 (2.8%)	5 (0.0%)	84693 (1.5%)	90376 (1.7%)	176 (0.0%)
Peripheral vascular disease	846 (0.0%)	31927 (0.4%)	67751 (1.4%)	32681 (0.6%)	10 (0.0%)	21523 (0.4%)	63290 (1.2%)	154 (0.0%)
VTE	188324 (1.6%)	154891 (1.7%)	128347 (2.6%)	79636 (1.4%)	6 (0.0%)	13712 (0.2%)	52607 (1.0%)	15834 (0.1%)

The description of the percentage of population exposed to selected medicinal products classes during the 365 days look-back period is presented in Table 14. Cardiovascular drugs were most prevalent: 17 to 26% across data databases, except PEDIANET. Drugs to treat mental health diseases ranged from 8 to 11%, but were not included in the CPRD data instance, which explains the low prevalence of mental health diseases in this data sources as these drugs are used as proxies to detect them. Lipid lowering drugs reached 10.5% in SIDIAP and 14% in BIFAP_PC. This drug class was not captured in the current CPRD instance. Antibiotic prevalence was between 13% and 25% across data sources.

Table 14. Demographic characteristics of study population and prevalence of selected medicines at 1-12-2020 (look back period 365 days)

	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
Study population	11827601 (100%)	9038207 (100%)	4880104 (100%)	5864586 (100%)	36587 (100%)	5746218 (100%)	5328274 (100%)	12237133 (100%)
Female	6139288 (52%)	4728817 (52%)	2485964 (51%)	2968662 (51%)	17710 (48%)	2891537 (50%)	2641540 (50%)	6070795 (50%)
Analgesics	2888260 (24.4%)	2222702 (24.6%)	1652496 (33.9%)	1405926 (24.0%)	532 (1.5%)	1281848 (22.3%)	971042 (18.2%)	0 (0.0%)
Antibiotics	1674103 (14.2%)	1197093 (13.2%)	1235627 (25.3%)	965448 (16.5%)	7997 (21.9%)	1167743 (20.3%)	896064 (16.8%)	2251916 (18.4%)
Antimycotics	53316 (0.5%)	37390 (0.4%)	0 (0.0%)	27971 (0.5%)	14 (0.0%)	90312 (1.6%)	0 (0.0%)	0 (0.0%)
Antithrombotic agents	1023568 (8.7%)	783968 (8.7%)	457731 (9.4%)	473538 (8.1%)	17 (0.0%)	538172 (9.4%)	544578 (10.2%)	573354 (4.7%)
Antivirals	55945 (0.5%)	44195 (0.5%)	55413 (1.1%)	32603 (0.6%)	95 (0.3%)	113830 (2.0%)	62462 (1.2%)	2495 (0.0%)
Systemic corticosteroids	397925 (3.4%)	272347 (3.0%)	252828 (5.2%)	235883 (4.0%)	1694 (4.6%)	143346 (2.5%)	221831 (4.2%)	71915 (0.6%)
Oncologic drugs	82198 (0.7%)	63518 (0.7%)	49361 (1.0%)	36925 (0.6%)	14 (0.0%)	8262 (0.1%)	43356 (0.8%)	29886 (0.2%)
Cardiovascular medication	2875552 (24.3%)	2168579 (24.0%)	1248131 (25.6%)	1252208 (21.4%)	91 (0.2%)	1435220 (25.0%)	1177032 (22.1%)	2128757 (17.4%)
Diabetes medications	719442 (6.1%)	527902 (5.8%)	350776 (7.2%)	341992 (5.8%)	40 (0.1%)	267846 (4.7%)	215528 (4.0%)	539146 (4.4%)

	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	IT-PEDIANET	DK-DHR	NO-NHR	UK-CPRD
Drugs to treat mental health diseases	1202246 (10.2%)	885846 (9.8%)	537303 (11.0%)	565074 (9.6%)	43 (0.1%)	482219 (8.4%)	432902 (8.1%)	0 (0.0%)
Obesity medications	594 (0.0%)	564 (0.0%)	56 (0.0%)	n/a	0 (0.0%)	5594 (0.1%)	13752 (0.3%)	0 (0.0%)
Antiepileptic drugs	444091 (3.8%)	324648 (3.6%)	230990 (4.7%)	223439 (3.8%)	127 (0.3%)	88066 (1.5%)	144619 (2.7%)	0 (0.0%)
Lipid lowering drugs	1659806 (14.0%)	1240614 (13.7%)	734823 (15.1%)	617787 (10.5%)	<5 (~%)	694684 (12.1%)	603127 (11.3%)	0 (0.0%)
Antiinflamatory and antirheumatic products	2138394 (18.1%)	1565446 (17.3%)	1375266 (28.2%)	1081734 (18.4%)	786 (2.1%)	687098 (12.0%)	813140 (15.3%)	0 (0.0%)
Sex hormones	336869 (2.8%)	256553 (2.8%)	123120 (2.5%)	121573 (2.1%)	18 (0.0%)	597299 (10.4%)	0 (0.0%)	0 (0.0%)
Immunosuppressive/immunomodulating agents	445533 (3.8%)	310871 (3.4%)	286645 (5.9%)	254366 (4.3%)	1797 (4.9%)	168581 (2.9%)	273526 (5.1%)	104651 (0.9%)

10.1.3 VACCINE COVERAGE

Below, we describe coverage data against a benchmarking reference for the corresponding country (red line), usually the latest available indicator reported in the Global Health Observatory by the World Health Organization (WHO) (29). For COVID-19 vaccines, the reference indicator was taken from the ECDC COVID-19 tracker (30). In few cases, we could not identify a reference indicator. In **Annex 3**, we present figures of vaccine coverage for all vaccine doses and the number of administrations.

10.1.3.1 Measles-containing vaccines

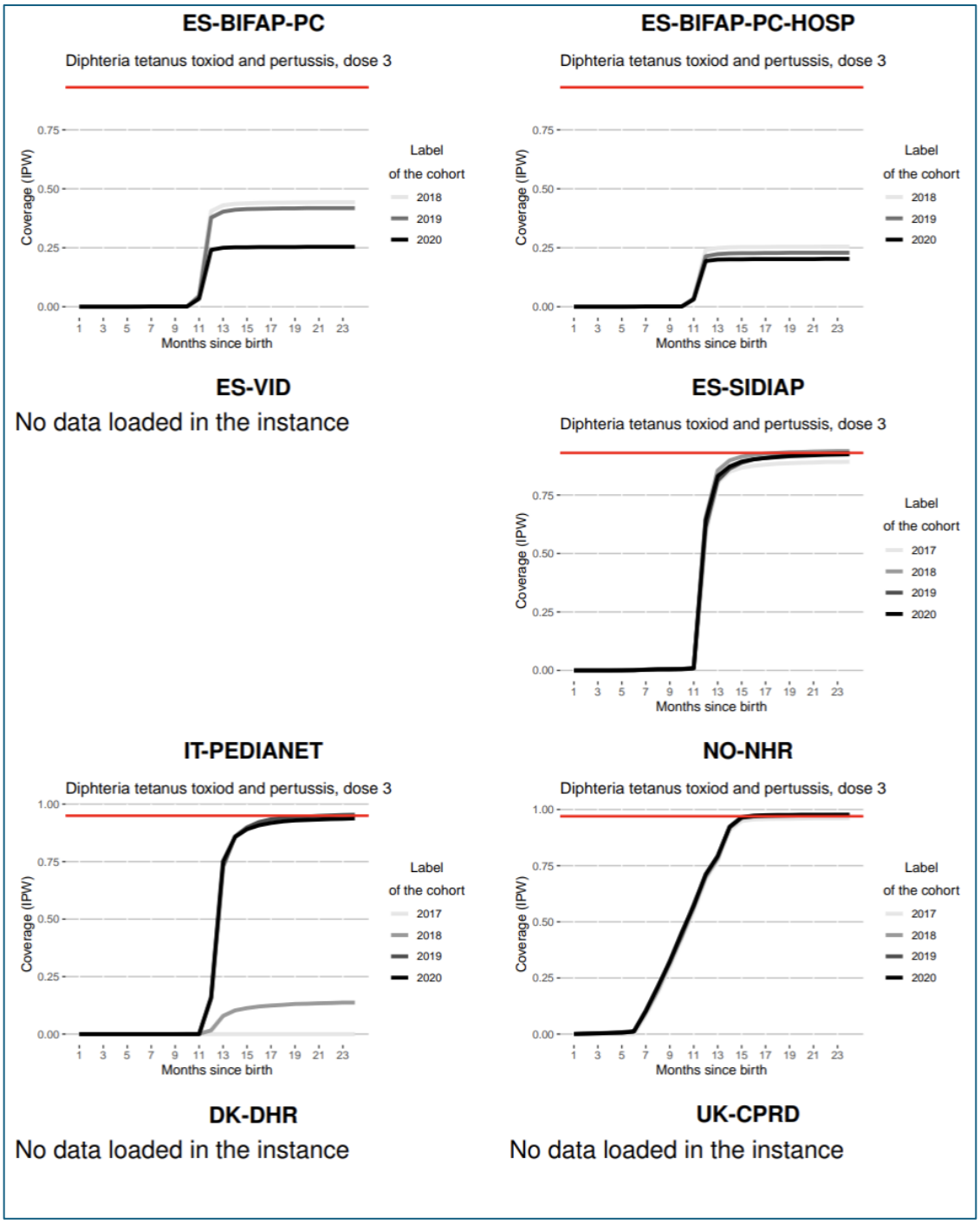
Figure 5 shows the coverage of the first dose of measles-containing vaccine in the birth cohorts 2017 to 2020, 2018, and 2019 in each data source compared to the WHO reference indicator (MCV1). BIFAP_PC and the PC-HOSP subpopulation overlap with the WHO reference values, same as SIDIAP and NHR. PEDIANET shows discordant coverage values against the reference line in the birth cohorts 2017 and 2018, whereas in the birth cohort 2019 the distance from the benchmarking is less than 10%. VID, DHR, and CPRD did not have data for this vaccination in the available data instance for this report.

Figure 5. Benchmarking of Measles-containing vaccines, first dose, by month of age for different birth year cohorts, versus WHO reference indicator value indicator (MCV1).

10.1.3.2 Diphtheria-Pertussis-Tetanus

The Figure 6 below shows the coverage of the third dose of Diphtheria-Pertussis-Tetanus vaccine in each data instance compared to the WHO reference indicator (DTP3) for the birth cohorts 2017 to 2020. BIFAP_PC_HOSP and BIFAP_PC- show estimates that deviate more than 10% from reference values for the whole study period. SIDIAP and NHR have overlapping estimates with reference values for all the birth cohorts. In PEDIANET, the birth cohorts 2017 and 2018 deviate more than 10% from the reference value, the birth cohort 2019 and 2020 overlaps well with the benchmark. VID, DHR, and CPRD did not have data for this vaccination in the available data instance for this report.

Figure 6. Benchmarking of Diphtheria-Pertussis-Tetanus vaccine, third dose, by month of age for different birth year cohorts versus WHO reference indicator value (DTP3).

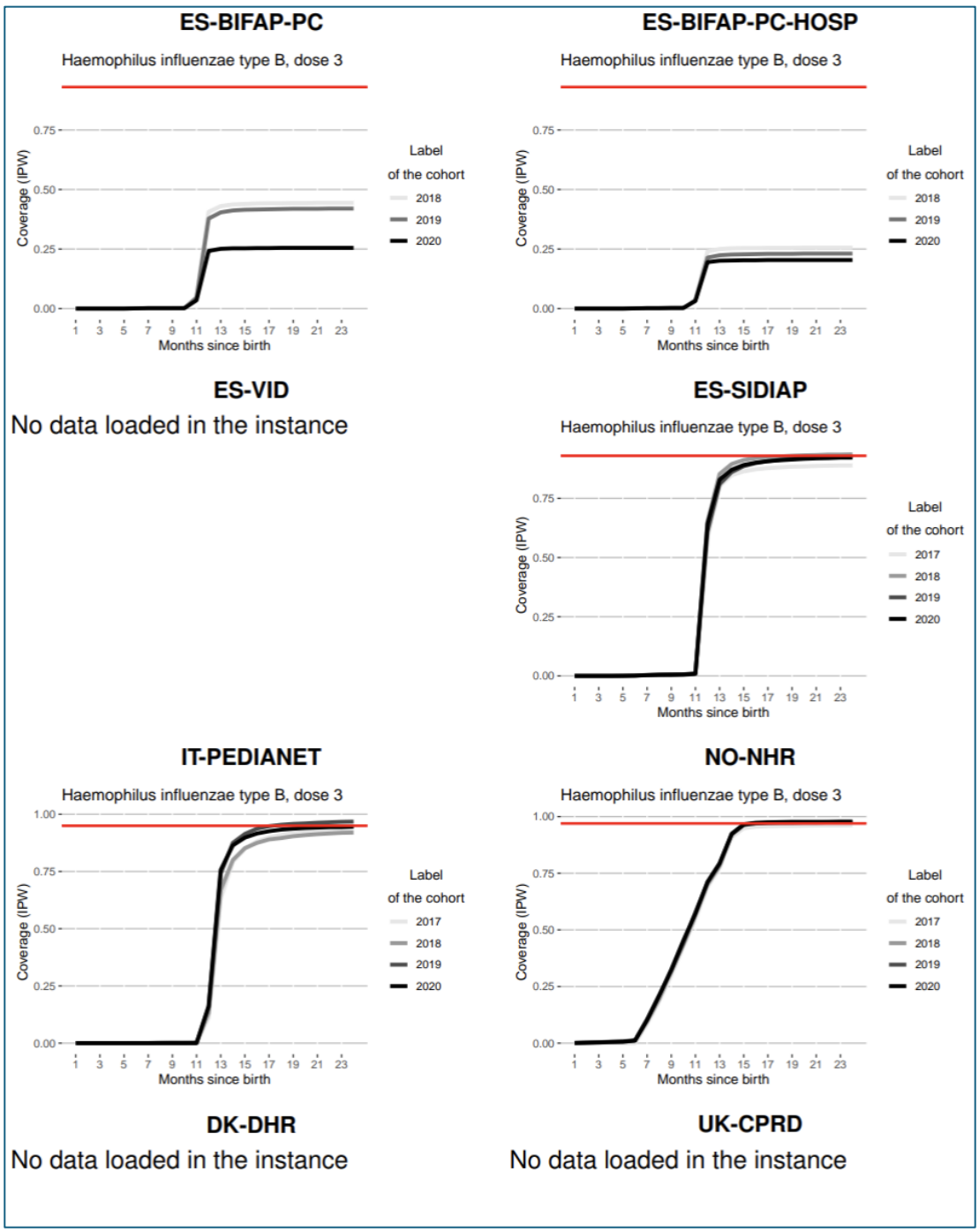


10.1.3.3 Haemophilus influenzae type B

Figure 7 shows the coverage of the third dose of *Haemophilus influenzae* type B vaccine in each data source compared to the WHO reference indicator (Hib3) in the birth cohorts 2017 to 2020.

BIFAP_PC and the PC-HOSP subpopulation show coverage values less than 50% and therefore far from the accepted 10% difference against the WHO reference value. SIDAP, PEDIANET and NHR have overlapping estimates with reference values for all the study period. VID, DHR, and CPRD did not have data for this vaccination in the available data instance for this report.

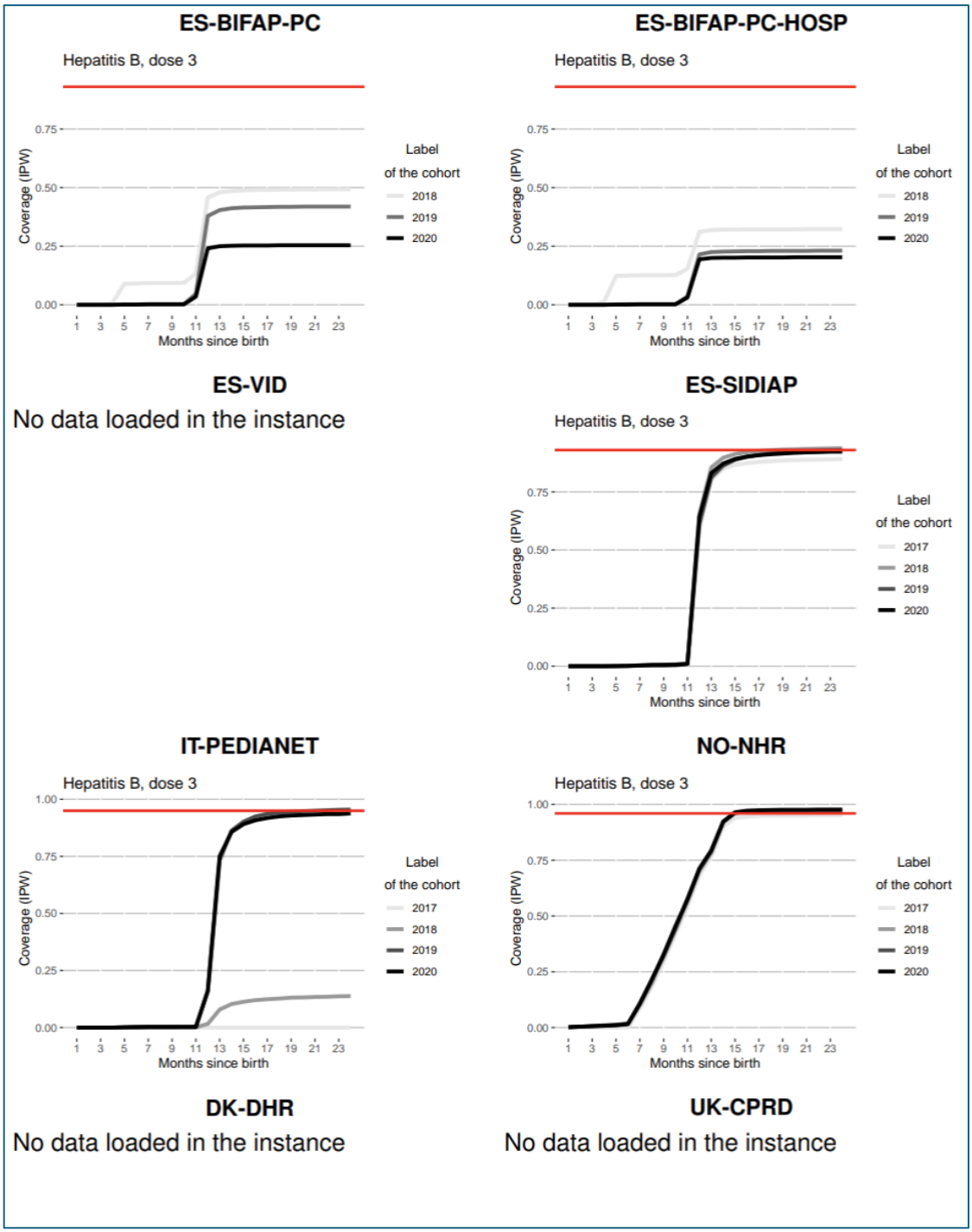
Figure 7. Benchmarking of Haemophilus influenzae type B vaccine, third dose, by month of age for different birth year cohorts versus WHO reference indicator value (Hib3).



10.1.3.4 Hepatitis B

Figure 8 shows the coverage of the third dose of Hepatitis B vaccine in each data source compared to the WHO reference indicator (HepB3) in the birth cohorts 2017 to 2020. Coverage values in BIFAP_PC and BIFAP_PC-HOSP are far below the WHO reference indicator value. Coverage values in SIDIAP and NHR overlap with the benchmark indicator, while PEDIANET fits well only for the birth cohorts 2019 and 2020. VID, DHR, and CPRD did not have data for this vaccination in the available data instance for this report.

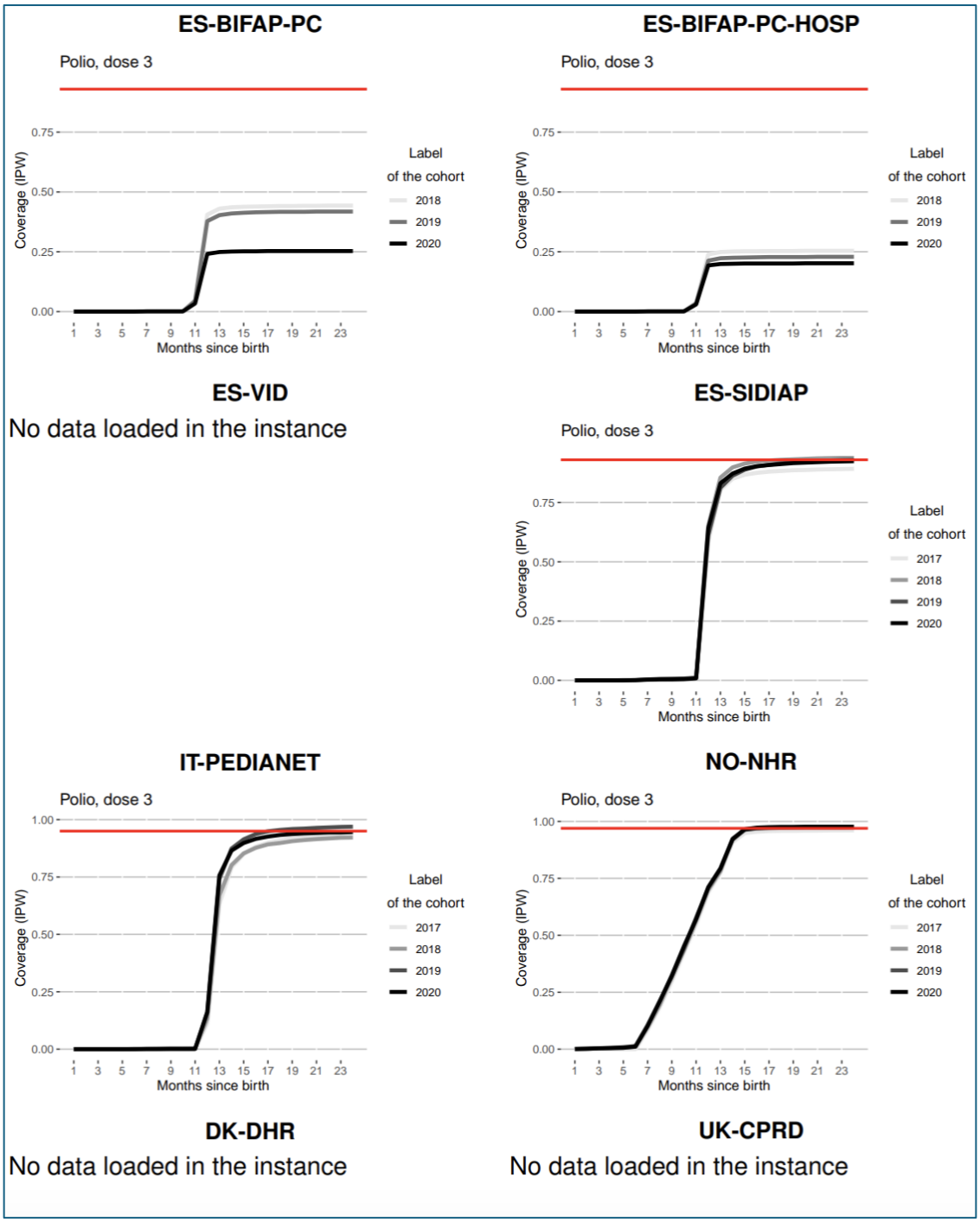
Figure 8. Benchmarking of Hepatitis B vaccine, third dose, by month of age for different birth year cohorts versus WHO reference indicator value (HepB3).



10.1.3.5 Poliomyelitis

Figure 9 shows the coverage of the third dose of Poliomyelitis vaccine in each data source compared to the WHO reference indicator (Pol3) in the birth cohorts 2017, 2018, and 2019 and 2020. SIDAP-ES, NHR-NO, and PEDIANET-IT have overlapping estimates with reference values in all birth cohorts. BIFAP-ES coverage values are more than 10% away from the reference WHO Spanish coverage. VID, DHR, and CPRD did not have data for this vaccination in the available data instance used in this report.

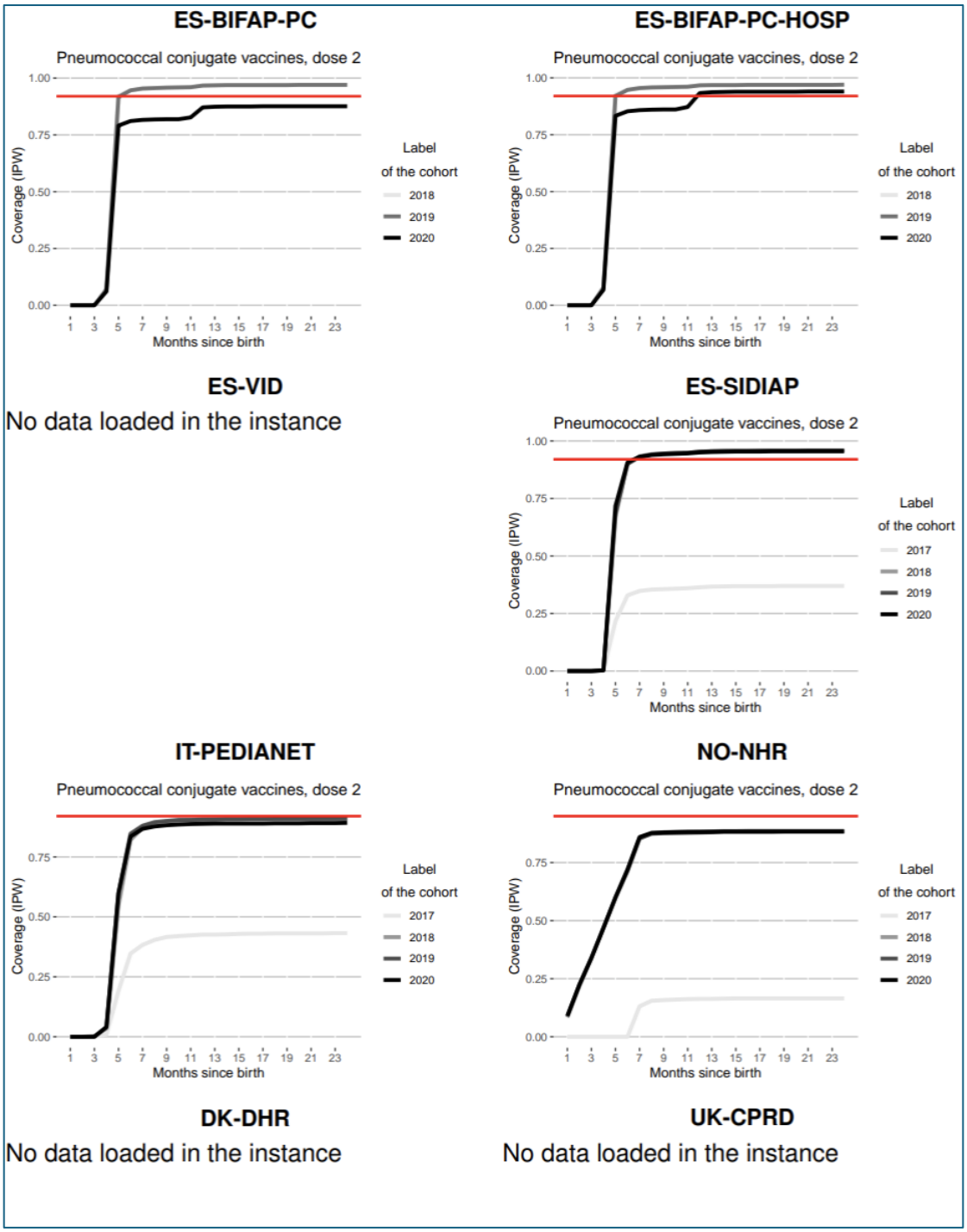
Figure 9. Benchmarking of Poliomyelitis vaccine, third dose, by month of age for different birth year cohorts versus WHO reference indicator value (Pol3).



10.1.3.6 Pneumococcal conjugate vaccines

Figure 10 shows the coverage of the second dose of Pneumococcal conjugate vaccine in each data source compared to the WHO reference indicator. VID, DHR, and CPRD do not have data for this vaccination in the data instance used for this report. Both BIFAP_PC_HOSP and BIFAP_PC show a close overlap with the benchmarking values for the birth cohorts 2018 and 2019, whereas 2017 deviate more than 10% from the reference value. SIDIAP has overlapping estimates with reference values in the 2018, 2019 and 2020. Values for the 2017 birth cohort are deviated by more than 10% from reference data. NHR and PEDIANET show values close to the reference data only for years 2018, 2019, and 2020, whereas values corresponding to the cohort of 2017 deviate more than 10% from benchmarking.

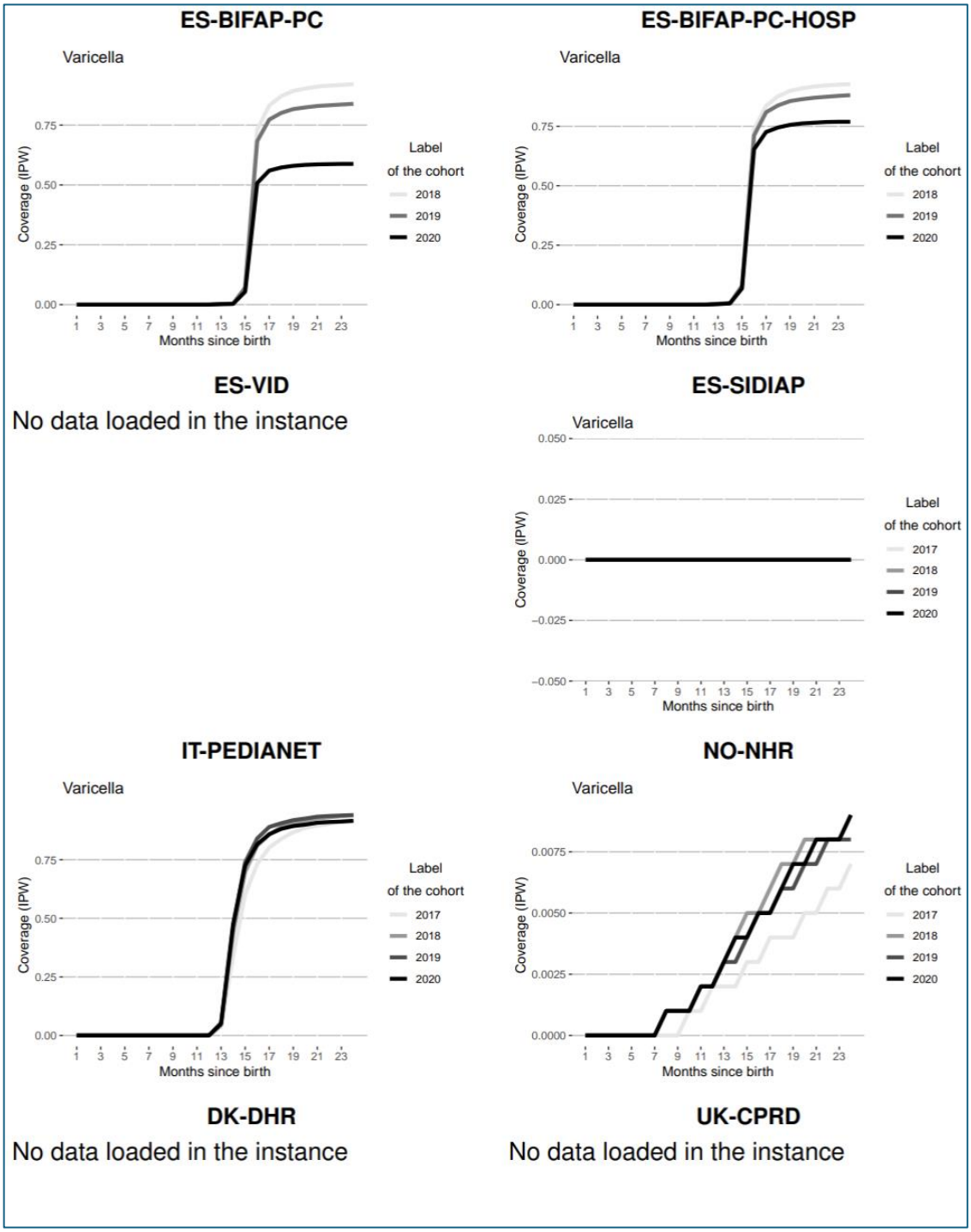
Figure 10. Benchmarking of Pneumococcal conjugate vaccine, second dose, versus WHO reference indicator value (PCV).



10.1.3.7 Varicella

No unique benchmarking source was available for varicella vaccination coverage. Figure 11 below presents the coverage proportions of varicella vaccine, 1 dose, in the birth cohorts 2017, 2018, 2019 and 2020. In PEDIANET-IT, the varicella coverage across birth cohorts reached more than 90%, which is close to the >75% reported in Veneto (31) and the 90% coverage reported in Puglia regions (32). In BIFAP-ES coverage crossed the 75% line in the BIFAP_PC cohort 2020 (33). VID, SIDIAP, DHR, and CPRD did not have data for this vaccination in the current data instance. Coverage values in NHR are very low.

Figure 11. Varicella vaccine coverage in SAFETY-VAC data sources.



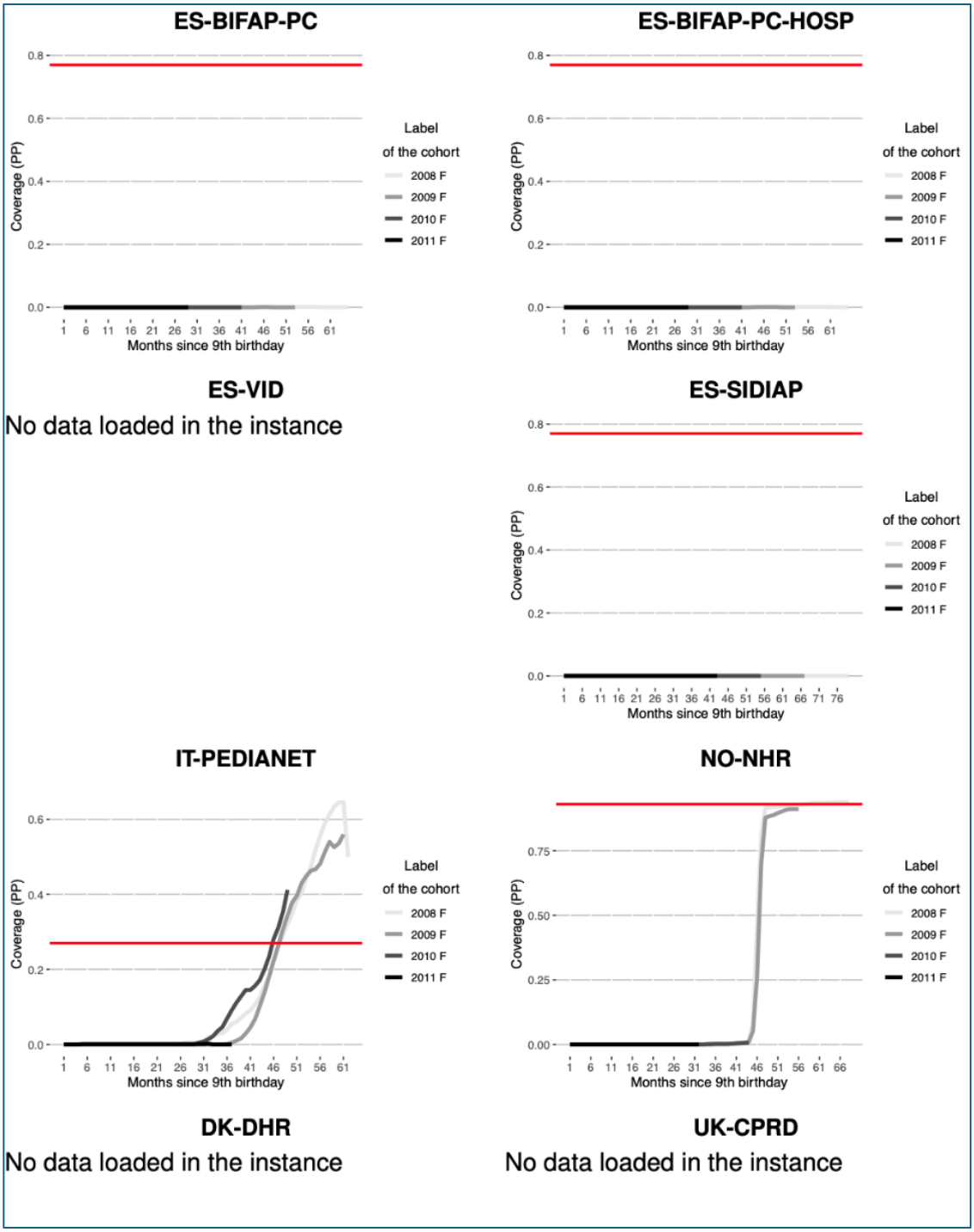
10.1.3.8 Bacille Calmette-Guérin (BCG)

BCG is a live-attenuated vaccine derived from *Mycobacterium bovis*. It is recommended to all healthy neonates in settings of high burden of tuberculosis. Nowadays, BCG vaccine is the most frequently administered of all vaccines worldwide (34). BCG vaccine is not recommended by any of the country authorities to which the databases of this study are coming from, therefore there is no information about this vaccine in this report, with exception of very few doses detected in BIFAP.

10.1.3.9 Human papillomavirus (HPV)

Figure 12 shows the coverage of the second dose of human papillomavirus vaccine in each data source compared to the WHO reference indicator (HPV) (2 doses in female adolescents). The cohort label refers to the birth year: study participants enter the cohort on the day of the 9th birthday and exit on the day of the 16th birthday. Only NHR and PEDIANET have available data for this analysis in the current data instance. NHR correlates with WHO reference values for the cohorts born in the 2008 and 2009. PEDIANET shows estimates that strongly deviate from the reference value, however it is close to the 73.3% coverage estimate calculated in the ADVANCE project for the Italian Val Padana database using the same coverage calculation method (PP.fu) (35).

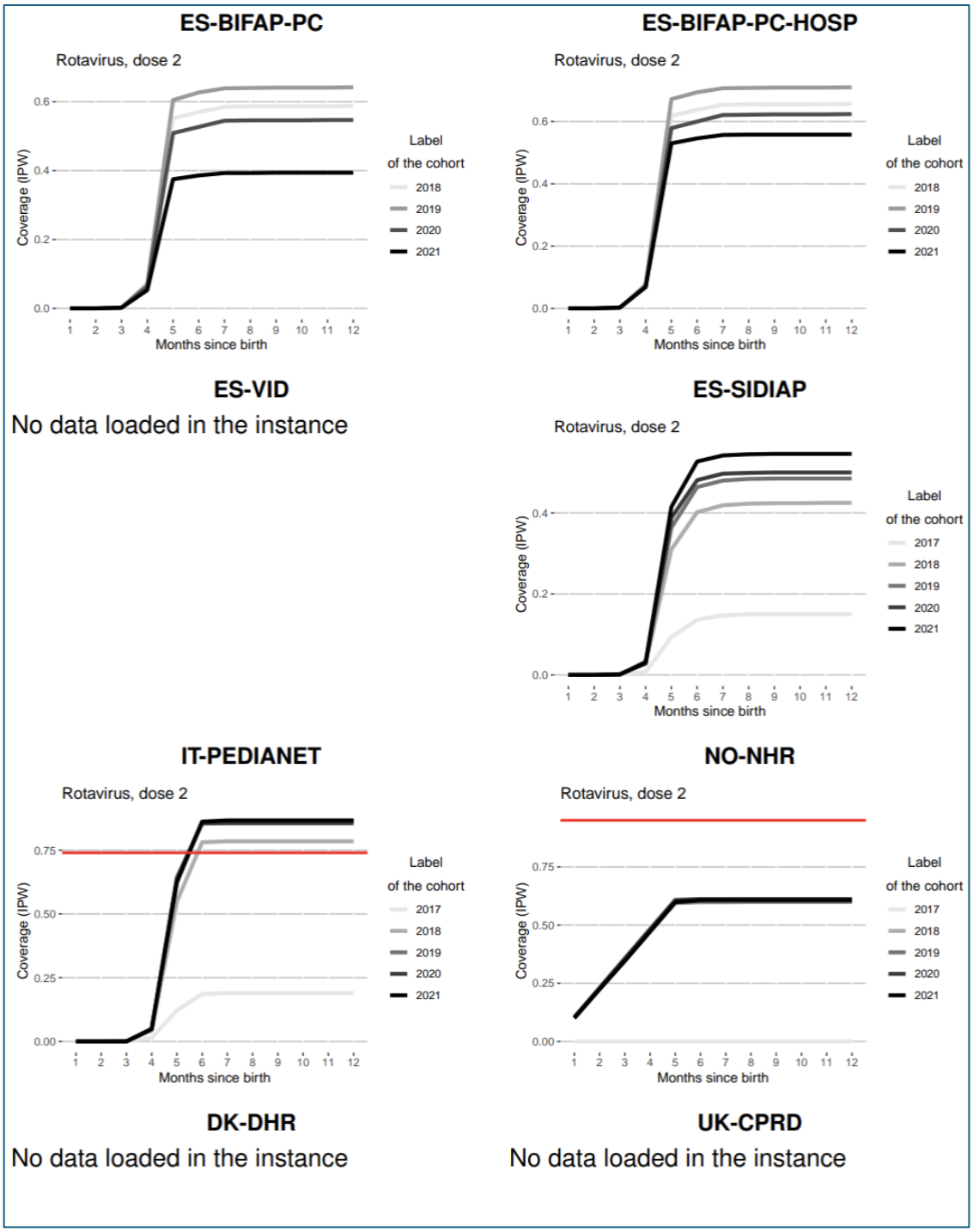
Figure 12. Coverage of human papillomavirus vaccine, second dose in female adolescents, and benchmarking versus WHO indicator value (HPV2).



10.1.3.10 Rotavirus

In the Italian PEDIANET database, the rotavirus vaccine coverage (2 doses) reached >70% across birth cohorts before the first year of age, except in the 2017 cohort. Figure 13 is aligned with the official coverage reported to the WHO (red line). NHR values deviate from benchmarking of more than 10% in all the study cohorts. In Spain, rotavirus vaccination is only funded by the National Health System for the premature babies; therefore, rotavirus vaccine coverage may differ substantially from region to region (there is no WHO reference value available for Spain). Ruiz-Contreras J., et al. (36) reports a national coverage ranging from 10 to 75%, aligned to estimates reported by BIFAP (several Spanish regions). Moreover, the same authors reported a coverage of 31-59% in Catalonia, which correlates well with the coverage reported in this study by the Catalonian SIDIAP database for the birth cohort 2018, 2019, 2020, and 2021.

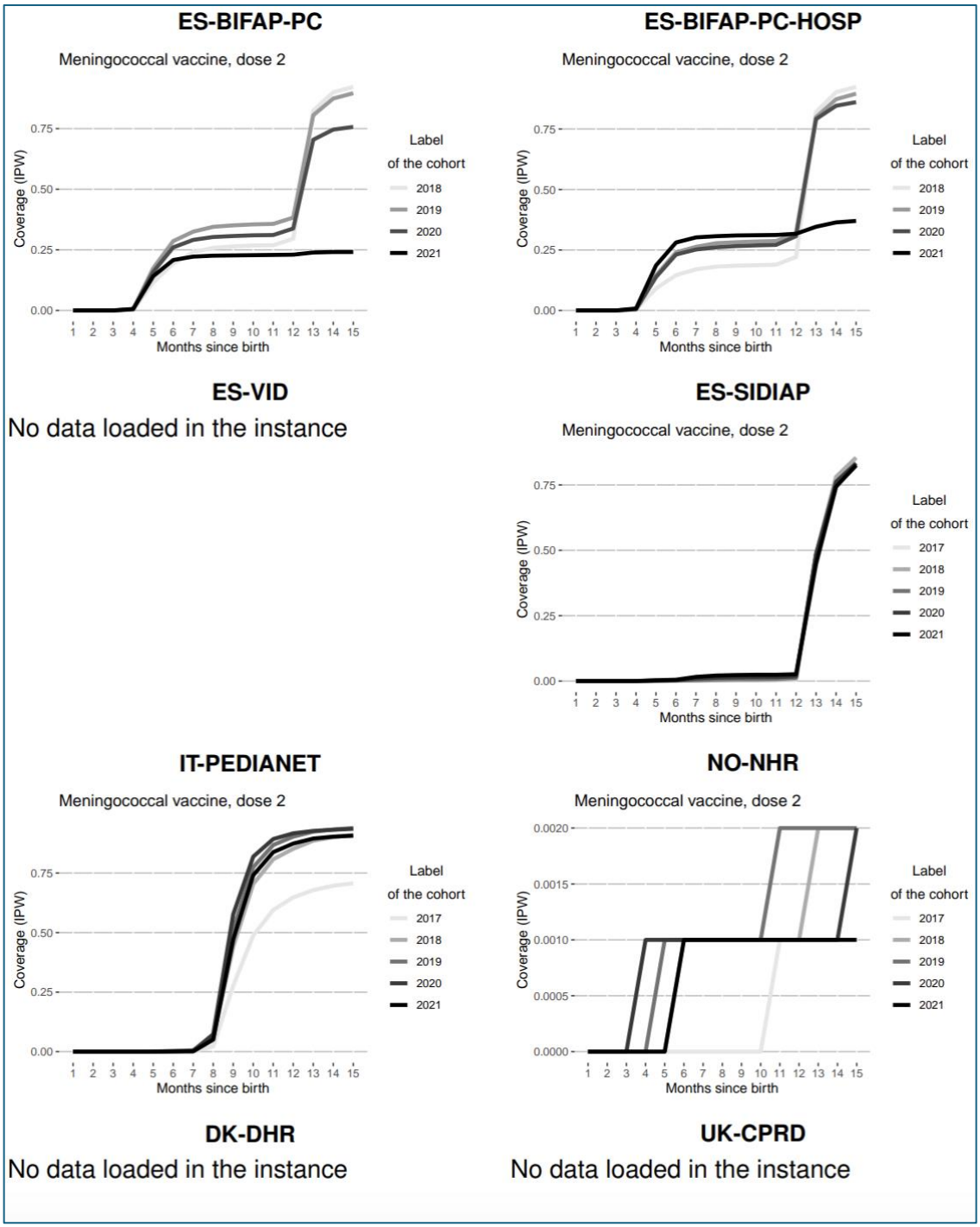
Figure 13. Benchmarking of rotavirus vaccine, second dose, by month of age for different birth year cohorts versus WHO reference indicator value (RotaC2)



10.1.3.11 Meningococcal vaccine

The meningococcal vaccine coverage reported annually through the WHO/UNICEF Joint Reporting Form on Immunization reported an administrative coverage of vaccine, at least one dose, between 68% to 89.5% in Spain in 2022 (37). Our coverage estimates in BIFAP and SIDIAP databases fit well within this range. In BIFAP, there is a lower coverage in the birth cohorts born in 2017 and 2021. In Italy, the WHO/UNICEF reports coverage figures between 55.4% to 80.9% in 2022, which is aligned to the PEDIANET coverage depicted in Figure 14. In NHR, the numbers are very low and required to be further explored. VID, DHR and CPRD did not presented data for this vaccine.

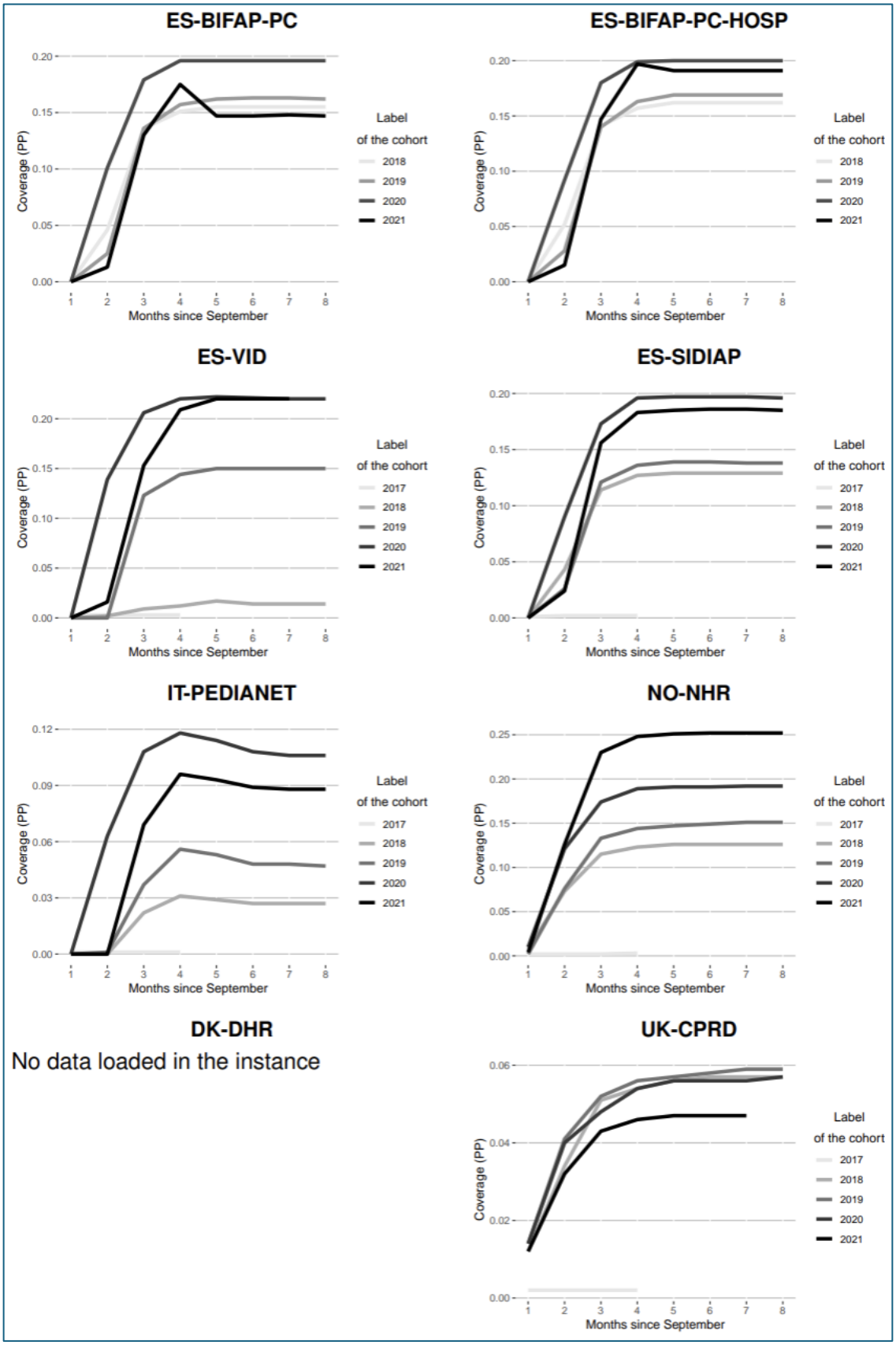
Figure 14. Second dose of meningococcal vaccine coverage by month of age for different birth year cohorts in SAFETY-VAC data sources.



10.1.3.12 Influenza vaccine

Except for DHR, data from all data sources are available and shown in the Figure 15. Coverage was calculated seasonally, each season starting in September and ending in April. All coverage values presented in this report fit well to the last influenza vaccine coverage report from the ECDC for the seasons 2018-2019, 2019-2020 and 2020-2021 (38). In general, our data shows an increase in the influenza vaccine coverage in the last two years of the temporal series, which corresponds well to the ECDC figures for Norway and Italy (there is no ECDC information for Spain and UK). In Spanish databases, the coverage figures are very consistent among them, they increased from about 15% in the seasons 2017, 2018, and 2019 to 20% in 2020 and 2021.

Figure 15 Influenza vaccine coverage in SAFETY-VAC data sources (all ages).

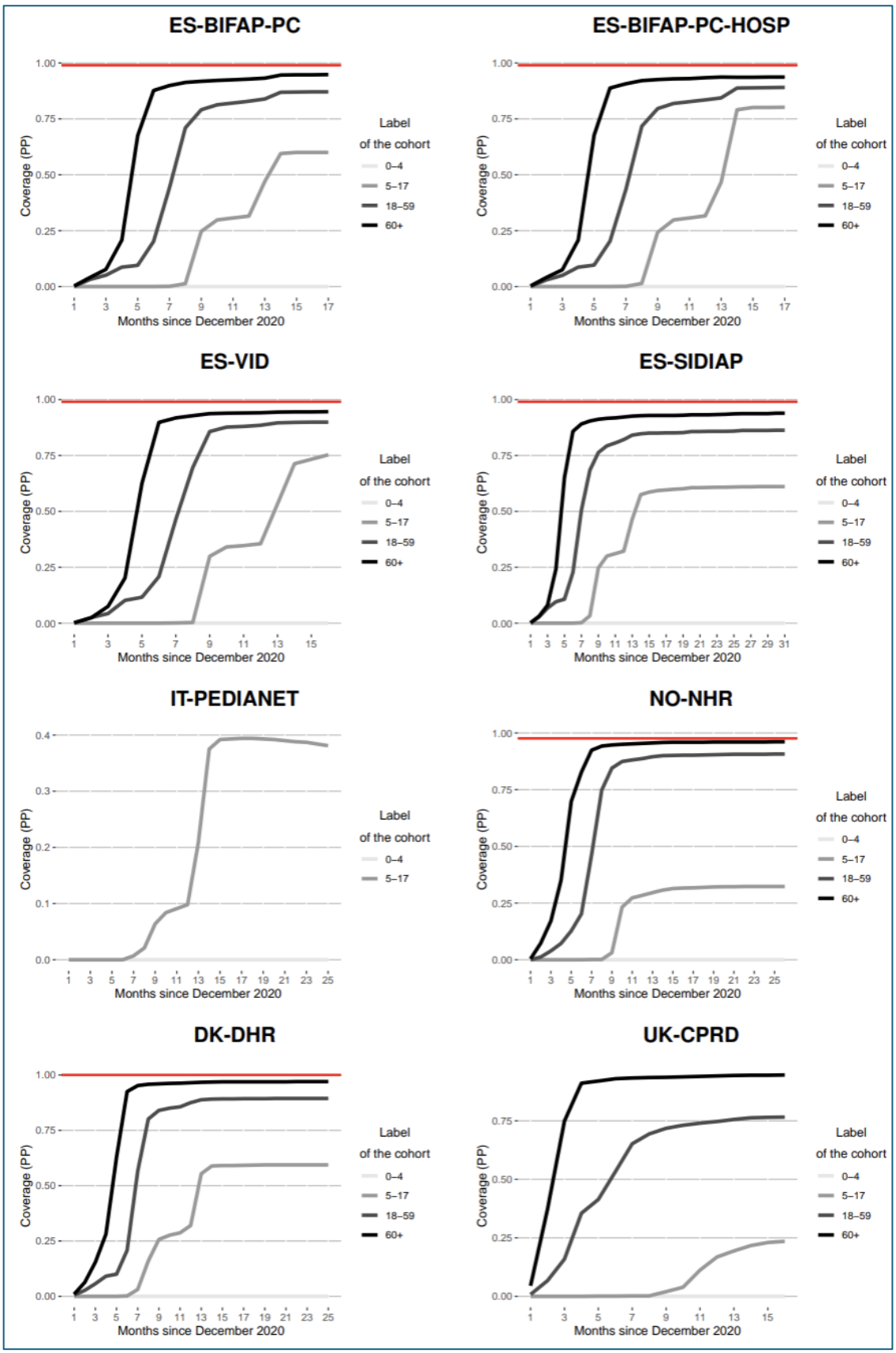


10.1.3.13 COVID-19 vaccines

Coverage of first to fourth dose of COVID-19 vaccines is presented in Figure 16 to Figure 19**.** In general, coverage figures fit well with ECDC reference value and with European and national recommendations of vaccination per age groups. The population cohort of > 60-year-old reached the highest coverage along the fourth doses, followed by adult population (18 to 59 yo). Recommendations to vaccinate adolescents with mRNA COVID-19 vaccines were launched between May and July 2021 (39,40) reaching, by the end of follow-up 75% and 50% coverage of dose 1 and 2 in the Spanish databases, respectively, 40% and 30% in Italy (PEDIANET), 60% and 50% in Denmark (DHR), 25% and 10% in Norway (NHR), and between 25% and 12% in the UK (CPRD). The coverage of the third dose in 18 to 59-year-old population dropped in all databases. Finally, in BIFAP, VID, CPRD and PEDIANET there was not enough follow-up time to report the coverage of the 4th dose. In SIDIAP, the coverage is approx. 7% below to the national benchmark (60%) and correlates well to the Catalonian data (approx. 10% difference) (41). In Denmark (DHR) and Norway (NHR), coverage figures in the population above 60-year-old correlates very well against the ECDC reference value. The coverage of the 4th dose in other age groups is low (max. 20%).

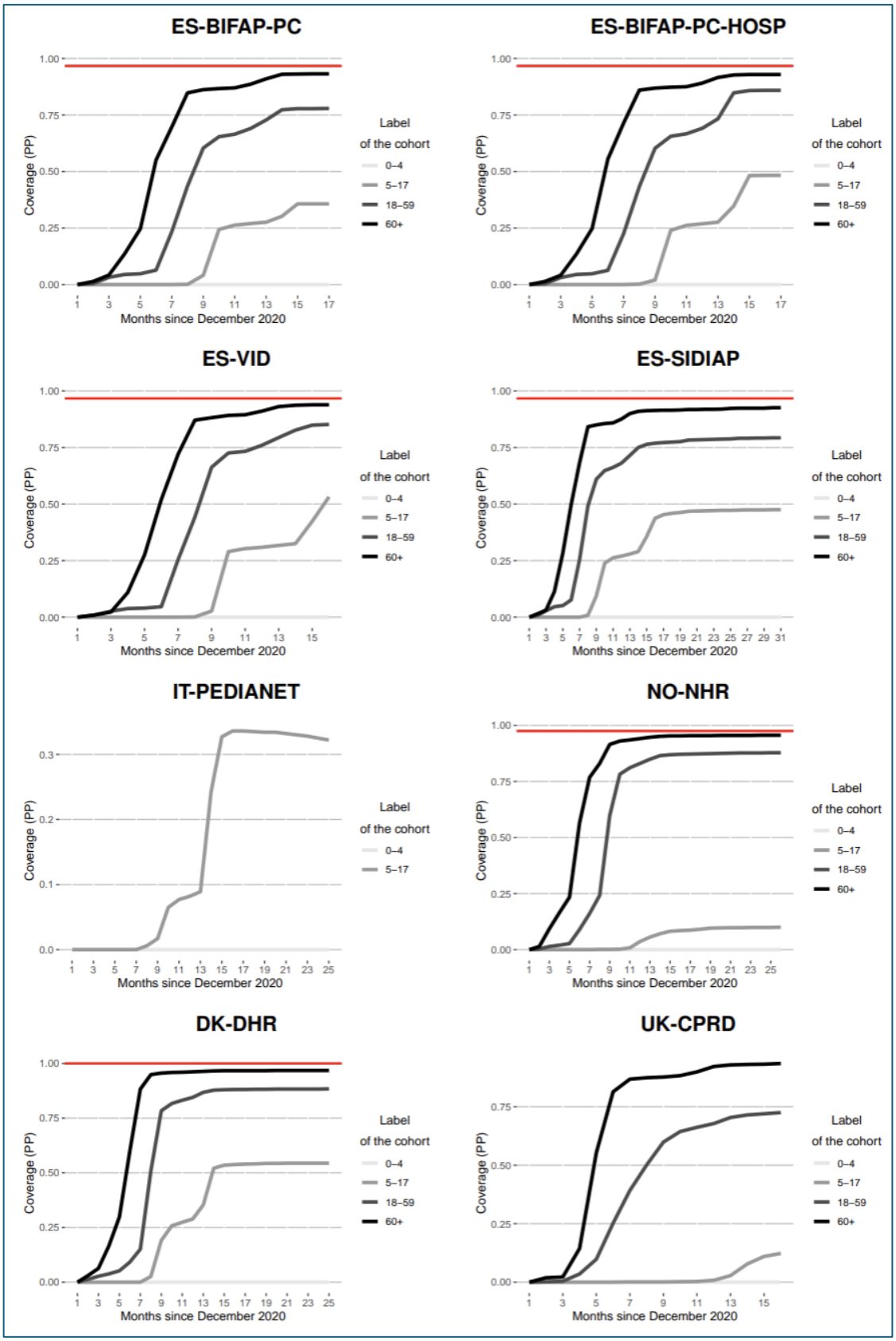
10.1.3.13.1 First Dose

Figure 16 Coverage of COVID-19 vaccine over calendar months (from 1/1/2020) first dose by age of population.



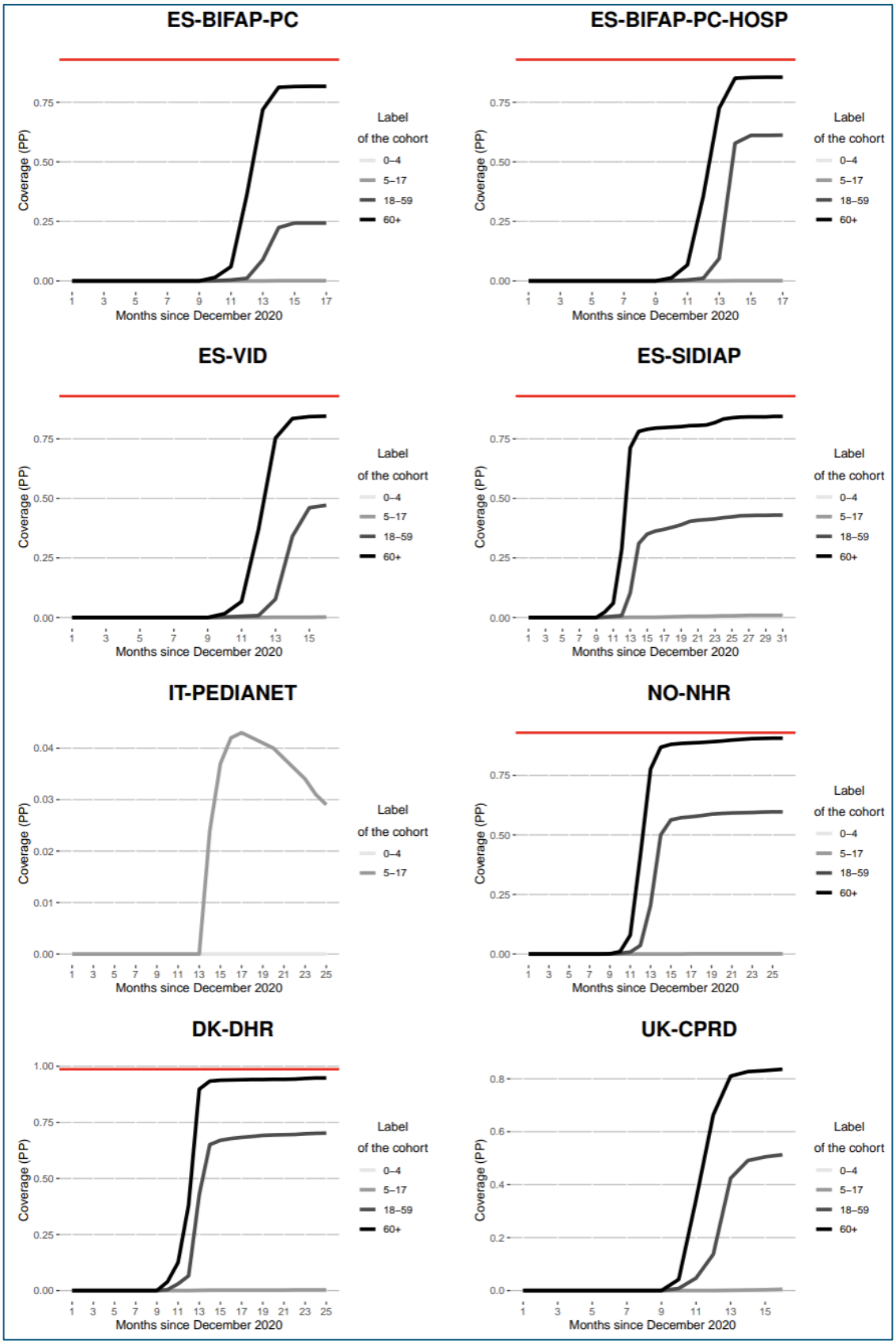
10.1.3.13.2 Second Dose COVID-19 vaccine

Figure 17 Coverage of COVID-19 vaccine over calendar months (from 1/1/2020) second dose by age of population



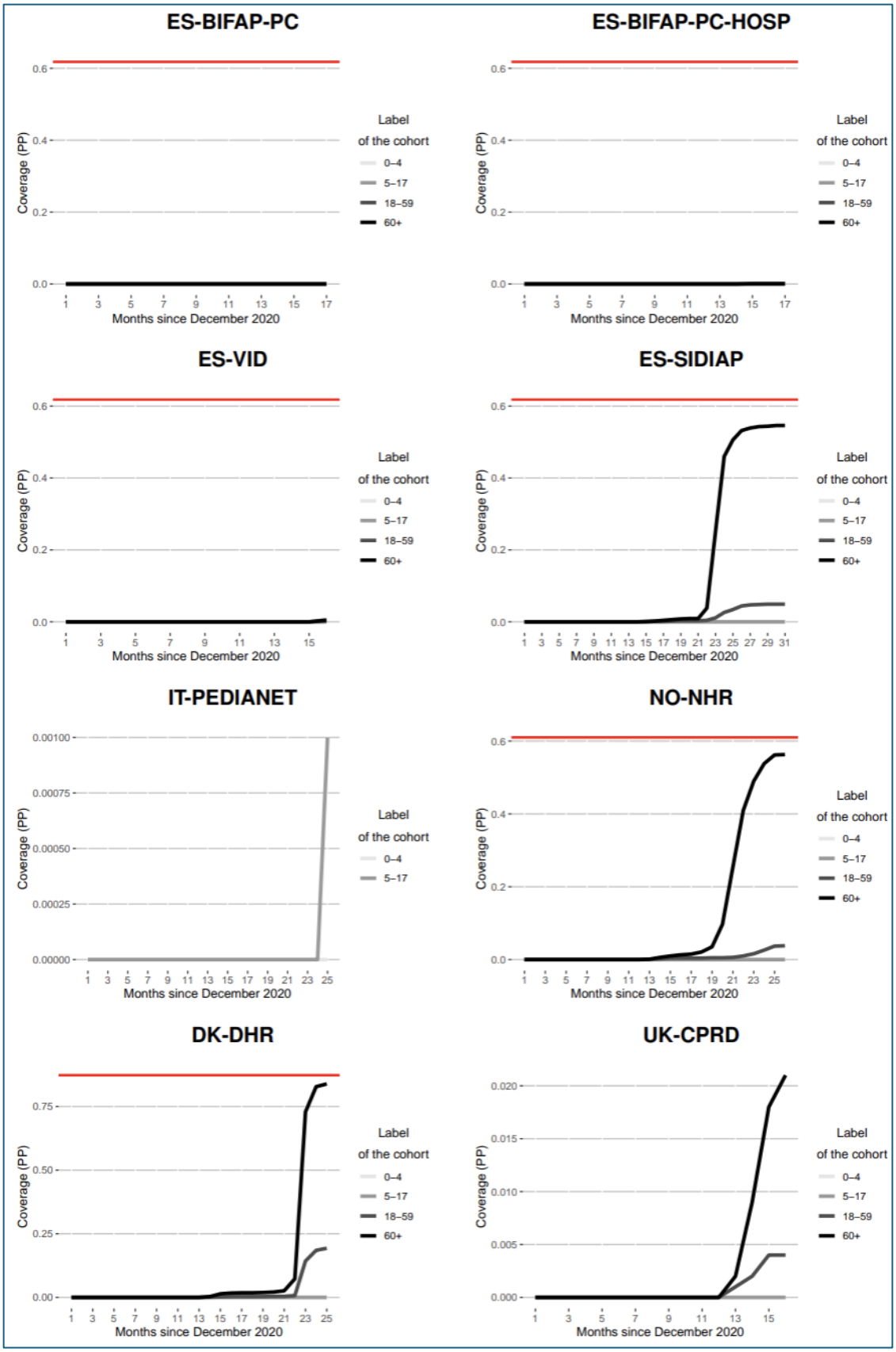
10.1.3.13.3 Third Dose COVID-19

Figure 18 Coverage of COVID-19 vaccine over calendar months (from 1/1/2020) third dose by age of population.



10.1.3.13.4 Fourth Dose

Figure 19 Coverage of COVID-19 vaccine over calendar months (from 1/1/2020) fourth dose by age of population.



10.1.4 INCIDENCE AND PREVALENCE OF STUDY OUTCOMES

This section reports the standardized incidence rates (IRs) per 100,000 person-year (PY) for 7 data sources from 5 countries. Standardization was performed using the official European demographic data. The observed incidence is compared to other published evidence for each event, where possible.
SAFETY-VAC data sources yearly standardized IRs are reported for each event in this section as well as in **Annex 4**, with 95% confidence intervals (CI). **Annex 5** shows age-specific IRs in 2019, whereas **Annex 6** shows standardised point and period prevalences for each event and data source. Specific code counts for each event during follow-up time and for all data sources are shown in **Annex 7**.

To better understand the observed rates from SAFETY-VAC data sources, it is important to keep in mind the overall nature of the data that is used to create the results, which are summarized below:

- UK-CPRD only includes data from primary care (≈100%)

- IT-PEDIANET mainly includes data from primary care (≈97%)

- ES-BIFAP_PC mainly includes data from primary care (≈78%) and dispensing from community pharmacy (≈18%)

- ES-VID mainly includes data from primary care (≈81%), and hospitalization and emergency room (≈13%) as well as intensive care unit (ICU) (≈1.9%).

- ES-BIFAP_PC_HOSP mainly includes data from primary care (≈60%), hospitals (≈25%, only related to the main hospitalization diagnosis), and dispensing from community pharmacy (≈12%)

- ES-SIDIAP mainly includes data from hospitals (≈48%), primary care (≈33%), dispensing from community pharmacy (≈6%) and ICU (≈5%).

- NO-NHR mainly includes data from hospitals and outpatients (≈48%), secondary care (≈32%) and primary care (≈11%)

- DK-DHR mainly includes data from hospitalization (includes specialist visits) and emergency room (>90% approximately).

This data composition is specifically associated with the current selection of the events of interest and the code list used to identify those events. Any changes in the list of the selected events and related codes may change the contribution of the different data provenance components.

10.1.4.1 Acute coronary artery disease (CAD)

CAD is a pathological process of coronary arteries that may derive from congenital abnormalities, atherosclerotic, or non-atherosclerotic causes (42,43) leading to inadequate blood supply to the myocardium. Acute symptoms include unstable angina and myocardial infarction. In this study, we defined the event as the combination of two main diagnoses: angina and acute myocardial infarction.

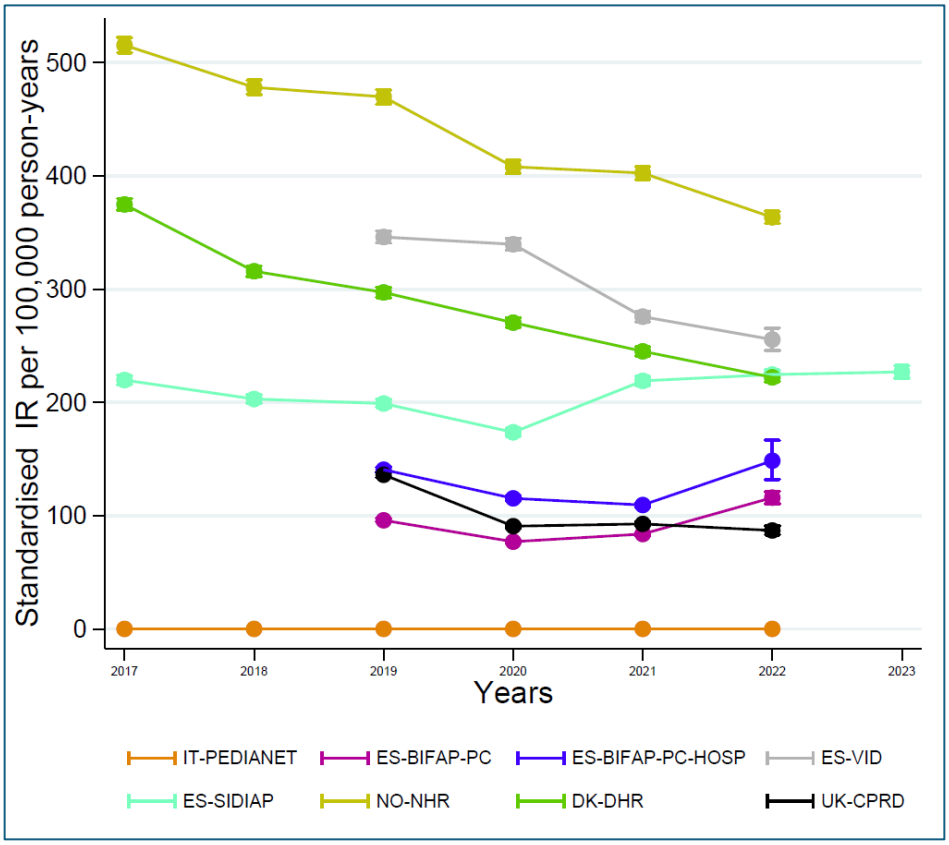
The SAFETY-VAC incidence rates of CAD are similar to benchmark IRs in Denmark but higher than national statistics in Norway.[4] UK-CPRD rates are lower than those previously reported in CPRD GOLD based on 2006 to 2015 data (44), however, unlike the current study, CPRD were linked to the hospital episode statistics, thus the difference is in the expected direction. Due to the nature of the event, rates are higher in databases containing diagnosis codes retrieved in hospital (VID, SIDIAP, NHR, DHR). BIFAP-PC-HOSP contain records from primary care linked to hospital records, however it only retrieves primary diagnoses while others identify primary and secondary diagnosis; this might produce lower rates in BIFAP_PC-HOSP. In general, CAD rates reported below are similar to the rates reported in the CVM study (22). There were no CAD cases identified in PEDIANET-IT due to the nature of the disease and age distribution of the study population.

Table 15 Standardized incidence rates of CAD per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				219.7	515.4	374.8	0.0
2018	0.0				203.0	478.2	315.8	214.4
2019	0.0	95.9	140.6	346.0	199.0	469.8	297.1	136.1
2020	0.0	77.0	115.3	339.6	173.6	408.0	270.4	90.7
2021	0.0	83.8	109.4	275.7	219.0	402.5	245.1	92.7
2022	0.0	115.9	148.5	255.5	224.6	363.4	222.1	86.9
2023					226.9			

[4] https://statistikkbank.fhi.no/hkr/

Figure 20 Standardized incidence rates of CAD per data source.



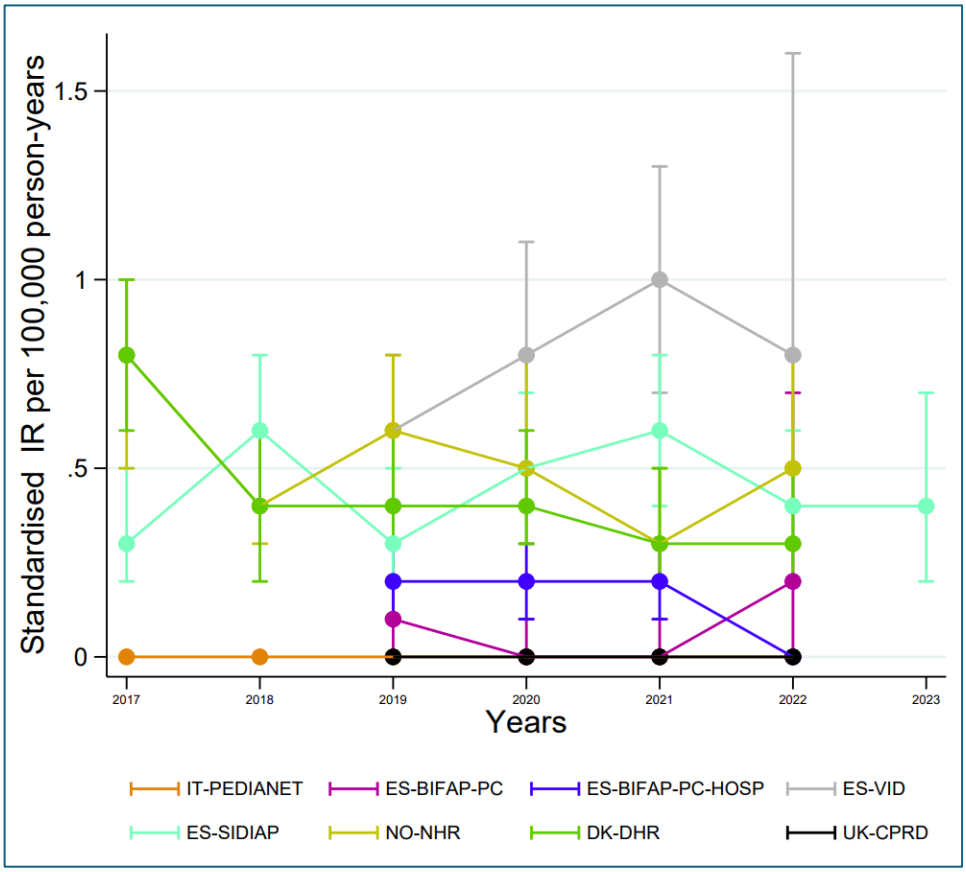
10.1.4.2 Acute disseminated encephalomyelitis (ADEM)

ADEM is an acute inflammatory autoimmune process of the central nervous system characterized by multiple foci of perivascular demyelination. It usually occurs shortly after an acute viral infection or immunization; however, it may coincide with the onset of infection or rarely no antecedent can be identified (45,46). The disease affects predominantly children, but it may also affect adults. In general, the estimated incidence of ADEM in childhood ranges from 0.2 to 0.8 per 100,000 per year (47). Reviews of non-interventional studies (NIS) from 2006 to 2014 report hospitalization rates of 0.5/100000 person year (PY) in the paediatric population (≤18 years old) with ADEM (negligible difference between males and females). Another study based in California from 1991 to 2000 reported ADEM rates of 0.4/100000 in < 20 years old, with a stratified incidence of 0.6 and 0.8/100,000 in 0-4 and 5–9 years old children (48). A prospective clinical study in Italy reported a similar incidence for ADEM of 1.1/100000 PY in individuals with <10 years of age (49). Similar ADEM rates were reported in the ADVANCE project for Spain and the UK, but slightly higher rates of 5.3/100000 PY were reported for the Italian data sources, as well as for the Denish data sources (between 2.0 to 7.1/100000 PY) (50). Herein, the generated annual IRs are in line with the overall values reported and observed in the literature and similar to the rates produced in the CVM project (22). In data sources having primary care and hospital data (SIDIAP, BIFAP_PC_HOSP, VID, NHR) rates are similar to the pooled analysis reported by Willame C., et al., in the ACCESS project (6) for this kind of data provenance: 0.33/100000 PY. For primary care-only (PEDIANET, BIFAP_PC) data sources, incidence rates are much lower, since this is not the setting where it is diagnosed and similar to the corresponding estimate in the ACCESS project: 0.05/100000 PY. No data was produced for this event for UK-CPRD.

Table 16. Standardized incidence rates of ADEM per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				0.3	0.8	0.8	0.0
2018	0.0				0.6	0.4	0.4	0.0
2019	0.0	0.1	0.2	0.6	0.3	0.6	0.4	0.0
2020	0.0	0.0	0.2	0.8	0.5	0.5	0.4	0.0
2021	0.0	0.0	0.2	1.0	0.6	0.3	0.3	0.0
2022	0.0	0.2	0.0	0.8	0.4	0.5	0.3	0.0
2023					0.4			

Figure 21. Standardized incidence rates of ADEM per data source.



10.1.4.3 Arrhythmia

Cardiac arrhythmias are defined as any disturbance of the normal beating of the heart or myocardial contraction. Classification of cardiac arrhythmias include the abnormalities in heart rate, disorders of electrical impulse generation, or impulse conduction. A prior clinical definition form is available (51).

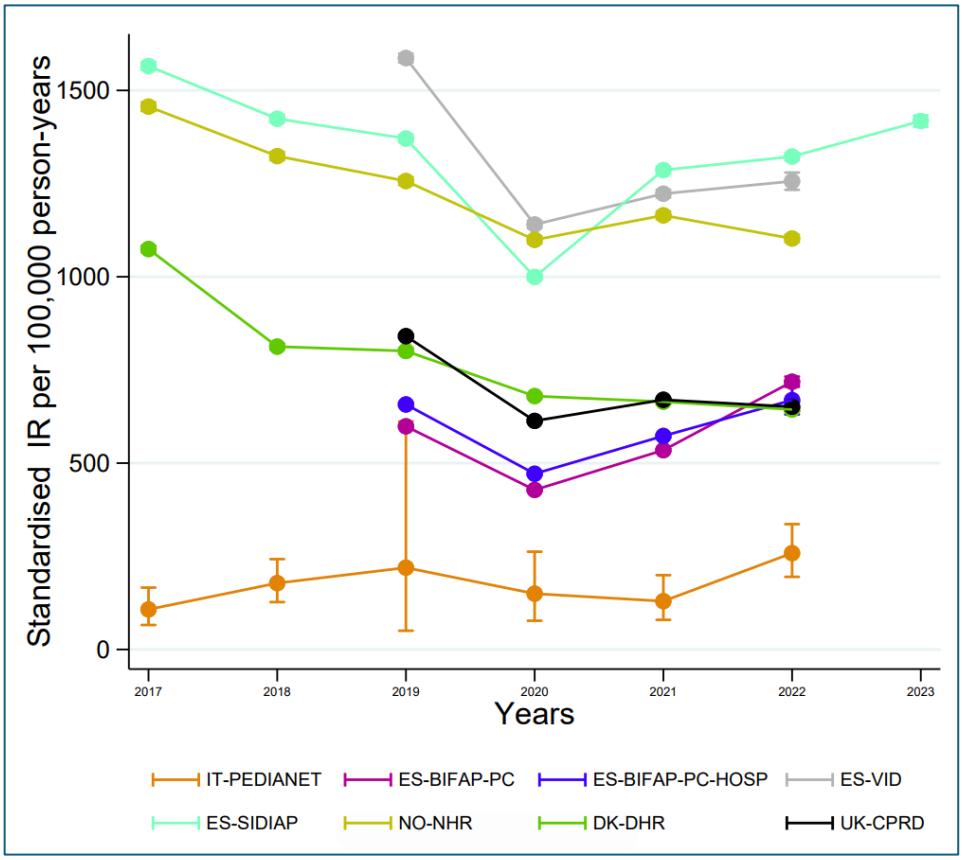
Atrial fibrillation (AF), the most common cardiac arrhythmia, affects 1.5-2% of European adults. The incidence of AF varies depending on race/ethnicity, with white individuals exhibiting a higher risk of AF when compared to black, asian, or hispanic individuals. Moreover, reported incidence widely differs across studies also due to methodological differences, such as case definitions (e.g., including valvular AF), or wash out period length, etc. Overall, in Western countries, population studies have reported AF incidence rates of 190 to 990/100000 PY(52,53).

In this assessment, the reported IRs from SAFETY-VAC data sources are in line with this range of values reported in western countries' literature. PEDIANET-IT shows lower ranges than the other data sources due to the inclusion of the paediatric population only.In databases combining GP and in-patient data, incidence rates are all close to the pooled rate produced in the ACCESS project (6) (1199.31/100.000 PY), except in BIFAP_PC_HOSP where rates ranged between 573 to 669/100000 PY, but this was similar to the 2019 rates produced in the CVM project (22). BIFAP does not include secondary discharge diagnosis. DHR values are in line with the above reported global annual incidence for this event, but a bit higher than other specific studies in Denmark that reported AF IRs varying from 367 to 481/100000 PY from 2004 to 2018 in adults probably due to the almost purely detection of this cases in the hospital and emergency room settings (54).

Table 17 Standardized incidence rates of arrhythmia per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	107.5				1565.5	1456.6	1074.4	0.0
2018	178.0				1424.1	1323.8	812.7	1130.7
2019	219.5	598.8	657.4	1586.9	1371.0	1256.9	800.9	840.7
2020	149.8	428.1	471.5	1140.3	999.5	1099.1	679.9	613.3
2021	129.6	534.7	572.9	1222.9	1286.3	1164.8	664.8	670.0
2022	258.5	718.3	669.1	1256.3	1322.6	1102.7	643.9	651.3
2023					1417.9			

Figure 22 Standardized incidence rates of arrhythmia per data source.



10.1.4.4 Arterial thrombosis

Arterial thrombosis is the pathological equivalent of the mechanism of hemostasis. Platelets adhere to collagen fibers surrounding the transected blood vessels, aggregate and form a plug. The initial trauma usually consists of damage to, or contraction of endothelial cells exposing subendothelial tissue to the blood stream. Arterial thrombosis frequently leads to rupture of the plaque in the artery wall and the ensuing thrombotic events are the triggers for acute ischemic injury in these diseases (55).

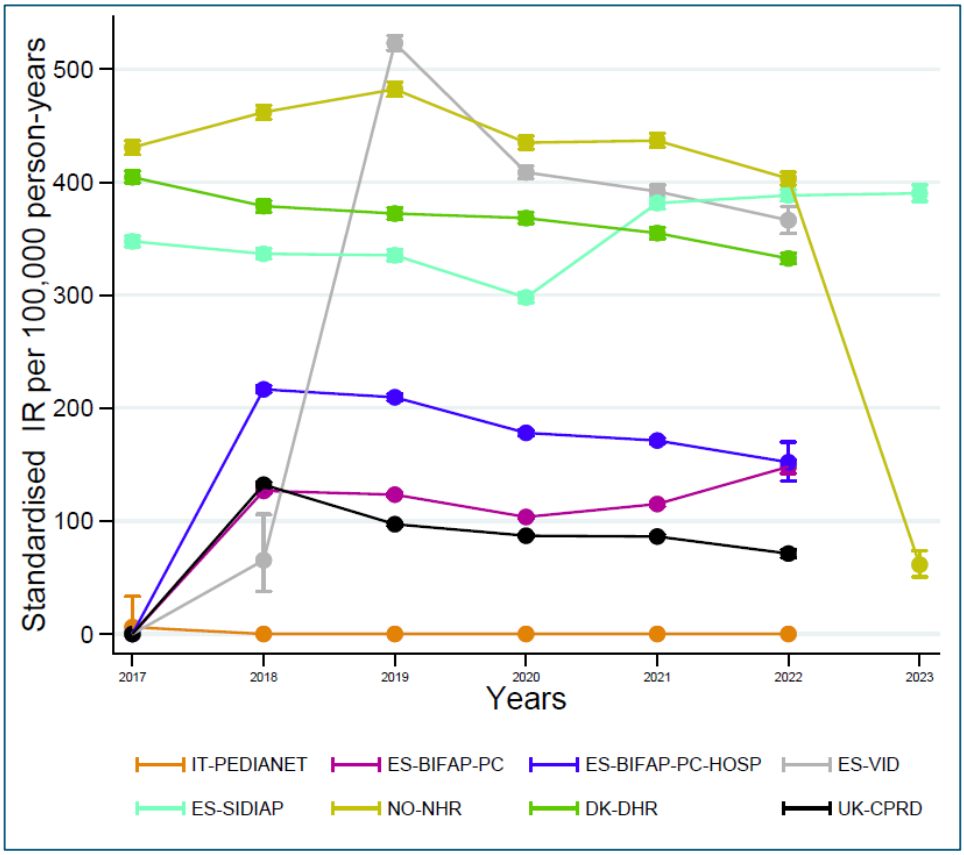
In this study, arterial thrombosis was defined as a combination of ischemic stroke and acute myocardial infarction. No published studies providing an overall estimation of ischemic stroke and/or acute myocardial infarction, or vice versa, have been reported so far. For instance, ischemic heart disease, also known as coronary artery disease, coronary heart disease, and atherosclerotic heart disease, is a group of diseases caused by atherothrombosis and its annual incidence is reported to be around 1500/100000 PY (2013 values) (56). A prospective registry study in Norway reported a standardized yearly incidence of peripheral artery disease (PAD) of around 290/100000 PY, but this study was performed in patients with hypercholesterolaemia only (57). Another Danish study reported PAD IRs ranging from 165 to 226/100000 PY from 2000–2018 (58).

PEDIANET rates are negligible due to the young population that this data source includes, which is not associated to the risk of this event. As expected, lower rates (<140/100000 PY) are observed in data sources reporting mainly data from primary care settings (CPRD and BIFAP_PC) compared to sources including primary care and/or hospitalization and emergency rooms data as this event is mainly managed in hospital and emergency settings.

Table 18 Standardized incidence rates of arterial thrombosis per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	6.0				347.7	430.9	404.5	0.0
2018	0.0				336.6	462.1	378.9	132.2
2019	0.0	123.4	209.6	523.2	335.4	482.4	372.3	97.2
2020	0.0	103.6	178.1	408.8	298.0	435.1	368.3	87.0
2021	0.0	115.1	171.3	391.9	381.6	436.9	355.0	86.3
2022	0.0	148.1	152.1	366.4	388.4	403.3	332.5	71.2
2023					390.2			

Figure 23 Standardized incidence rates of arterial thrombosis per data source



10.1.4.5 Autoimmune hepatitis

Autoimmune hepatitis (AIH) is a chronic, inflammatory liver disease characterized by circulating autoantibodies and elevated serum globulin levels. There are three AIH subtypes: AIH plus primary biliary cholangitis overlap, AIH plus primary sclerosing cholangitis overlap, and AIH with autoimmune cholangitis (59).

AIH varies significantly in incidence and prevalence across countries and regions. From a massive recent review of 37 studies that includes more than 239 million participants and 18 countries across five continents, the global pooled incidence of AIH was found to be 1.28/100,000 PY. The incidence of AIH was greater in countries with high Human Development Index (>0.92), among females and adults when compared with children (60). For instance, from a Norwegian search in patient databases and medical records from 1985-1994, the mean annual incidence was 1.6/100,000, slightly higher than the global values (61) and in line with another study from nationwide healthcare registries in Denmark, with a reported AIH IR of 1.68/100,000 PY (62).

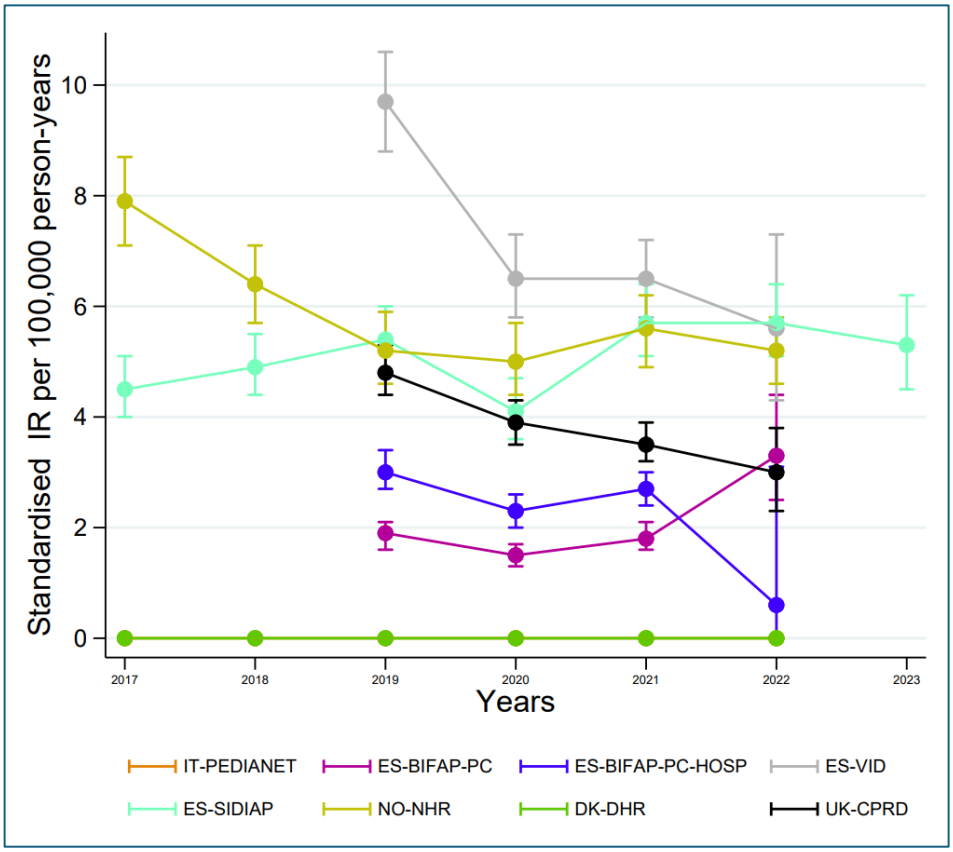
In general, incidence rates reported in this study by primary care only (BIFAP-PC, CPRD) and primary care-hospital data sources (BIFAP_PC_HOSP, VID, SIDIAP, NHR) are slightly higher (1.5 to 9.7/100.000 PY) than the rates reported previously in Europe: 0.9 to 2.0/100.000 PY (63).

There were no cases reported in PEDIANET as well as no cases in DHR due to information coming mainly from hospitalization and emergency room settings, where this event is not generally detected and managed. We could have been overestimating incidence rates due to the lack of specificity of the diagnosis codes to properly identify AIH cases, which ideally require histological confirmation, biochemical and autoimmune serological tests, and the assessment by a medical specialist (64).

Table 19 Standardized incidence rates of autoimmune hepatitis per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				4.5	7.9	0.0	0.0
2018	0.0				4.9	6.4	0.0	8.4
2019	0.0	1.9	3.0	9.7	5.4	5.2	0.0	4.8
2020	0.0	1.5	2.3	6.5	4.1	5.0	0.0	3.9
2021	0.0	1.8	2.7	6.5	5.7	5.6	0.0	3.5
2022	0.0	3.3	0.6	5.6	5.7	5.2	0.0	3.0
2023					5.3			

Figure 24. Standardized incidence rates of autoimmune hepatitis per data source.



10.1.4.6 Bell's palsy

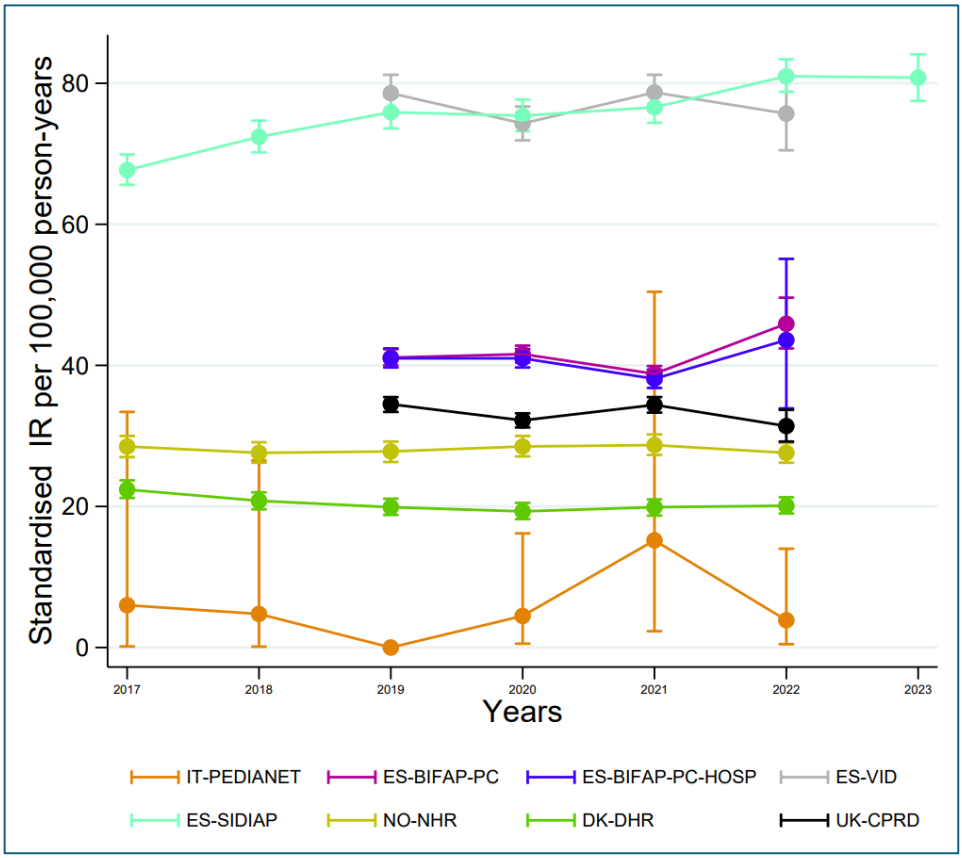
Bell's palsy is a peripheral neuropathy that causes facial nerve palsy; most cases are unilateral and occur unexpectedly, with rapid progression (worsening over a short period of time).

Depending on the geographical regions under study, different values for the annual incidence of Bell's palsy are reported, ranging from 13 to 107 cases per 100,000 PY (65–67). In general, the rates reported by SAFETY-VAC data sources are in line with the annual incidence reported in other studies. European nordic countries data sources (CPRD, DHR, NHR) reported lower rates than Spanish data sources (BIFAP-PC, BIFAP_PC-HOSP, VID, SIDIAP), as can be observed in the table and figure below. However, Spanish data are in line with previously reported rates in other studies (68). PEDIANET paediatric population reported lower rates than generally reported incidence in Europe. Incidence rates across data sources are very well aligned to the previous CVM (22) and ACCESS (6) projects. In PEDIANET-IT, there is a peak in the incidence of Bells' palsy in 2021 (15.2/100.000 PY), meaning three times the average annual incidence rate. In CPRD, rates from other evidence were between 20.9 and 30 per 100000 PY and were well aligned with those previously reported in studies unrelated to SAFETY-VAC data sources (69,70).

Table 20 Standardized incidence rates of Bells' palsy per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	6.0				67.7	28.5	22.4	0.0
2018	4.8				72.4	27.6	20.8	37.4
2019	0.0	41.1	41.0	78.6	75.9	27.8	19.9	34.5
2020	4.5	41.6	41.0	74.3	75.4	28.5	19.3	32.2
2021	15.2	38.8	38.1	78.7	76.6	28.7	19.9	34.4
2022	3.9	45.9	43.6	75.7	81.0	27.6	20.1	31.4
2023					80.8			

Figure 25 Standardized incidence rates of Bells' palsy per data source.



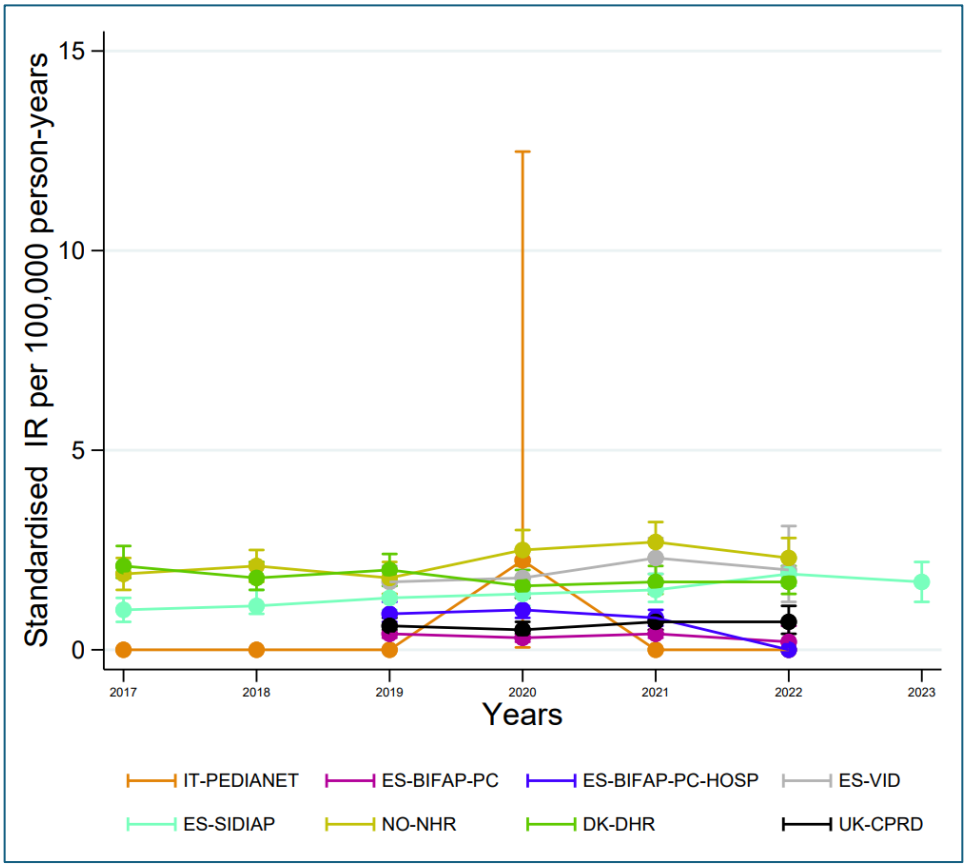
10.1.4.7 Cerebral venous sinus thrombosis (CVST)

Cerebral venous sinus thrombosis (CVST) is a stroke subtype. It occurs when a blood clot forms in the brain's venous sinuses. This prevents blood from draining out of the brain. As a consequence, blood cells may break into the brain tissues, leaking blood and forming a haemorrhage (71). CVST is a rare event with IR point estimates generally in the range of 0.21–1.72/100000 PY in 2015–2019 (70) and higher incidence in primary care-hospital linked databases (BIFAP_PC_HOSP, VID, SIDIAP, NHR) and manly hospital linked data sources (DHR) than in primary care-only data (CPRD, BIFAP_PC, PEDIANET). In BIFAP_PC_HOSP the incidences were close to 1/100000 PY, with no events detected in 2022, it should be noted that BIFAP_PC_HOSP data do not include secondary discharge diagnoses, in contrast to other data sources. BIFAP_PC_HOSP IRs in this study is similar to the rate reported in the CVM study: 0.92/100. 000 PY (2019) (22). In VID, rates ranged between 1.73 to 2.3/100000 PY, similar to 1.72/100000 PY detected in the CVM project in 2019. In SIDIAP, the average incidence rate was 1.41/100000, close to the 2019 CVM rate of 1.24/100000 PY. In Norway (NHR), the incidence rates are 1.9/100.000 PY (2017), 2.1 (2018), 1.8 (2019), 2.5 (2020), 2.7 (2021), and 2.3 (2022), which is higher than the rate of 1.0/100.000 PY reported by Pottegård et al. for the Norwegian data (72). DHR IRs are also similar to NHR, ranging from 1.6 to 2.1/100000 PY and higher than the rates reported by Pottegård et al. (72). In primary care-only databases (CPRD and BIFAP_PC) rates are comparable to the incidences reported in the CVM project (approx. 0.5/100000P Y) and previously published evidence (70). Finally, PEDIANET-IT reported cases in 2020 only, 2.2/100000 PY. Except for NHR, CVST rates in this study are aligned with pre-pandemic rates reported elsewhere (73).

Table 21 Standardized incidence rates of CVST per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				1.0	1.9	2.1	0.0
2018	0.0				1.1	2.1	1.8	0.8
2019	0.0	0.4	0.9	1.7	1.3	1.8	2.0	0.6
2020	2.2	0.3	1.0	1.8	1.4	2.5	1.6	0.5
2021	0.0	0.4	0.8	2.3	1.5	2.7	1.7	0.7
2022	0.0	0.2	0.0	2.0	1.9	2.3	1.7	0.7
2023					1.7			

Figure 26 Standardized incidence rates of CVST per data source.



10.1.4.8 Diabetes type 1

Diabetes mellitus groups some metabolic diseases defined by metabolic and hormonal changes that cause hyperglycemia and defects in insulin secretion. Type 1 diabetes results from b-cell destruction. It usually leads to absolute insulin deficiency and makes the patient become dependent of insulin from early ages of life (74).

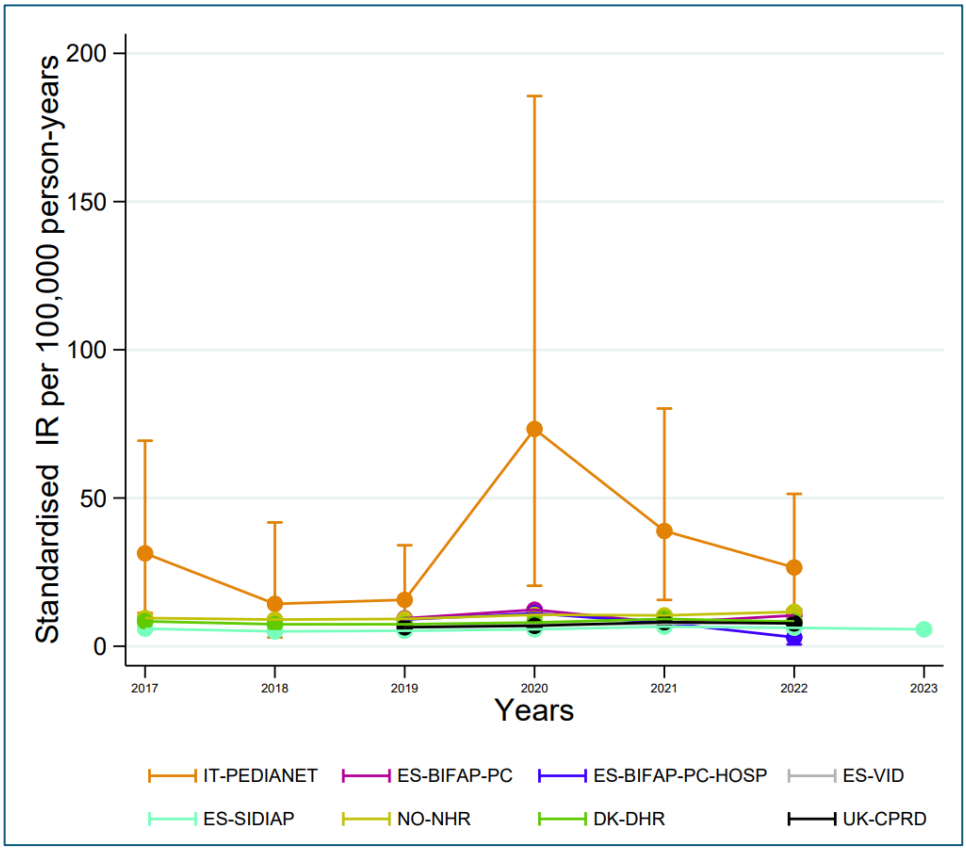
The results of a meta-analysis of 193 studies showed that the incidence of type 1 diabetes was 15/100,000 PY (75). (75)In a Denmark population-based cohort study including all liveborn infants who were born between 1980 and 2009, the total IR showed to be 17.7/100,000 PY. IRs of Diabetes type I increased with age until youth, raising from 4,40/100,000 PY in babies <1 year old up to 25,77/100,000 PY in the age range 10-17 (76). Another study from existing population-based healthcare registers in Denmark found similar IRs: 0.35/1000 PY in men and 0.25/1000 PY in women up to late adolescence (77), which correspond to the IRs found in other countries like Sweden, Finland, Norway and UK (IRS over 0.20/1000 PY) (78). Also, IRs in younger ages were slightly increasing, whereas rates in older ages showed a decrease; the overall average a decrease of 3.5% per year (77).

Due to the difficulties to capture diabetes type 1 through diagnosis codes only, in this study we identify diabetes type 1 using insulin and other antidiabetics dispensed/prescribed in childhood and in early years of adulthood (25 years of age) as a proxy of the event. We did not use diagnosis codes. The ACCESS project used a similar approach, therefore rates are comparable (6). In primary care-hospital linked databases (BIFAP_PC_HOSP, SIDIAP, VID, and NHR) rates are < 11.6/100.000 PY along the study period.. In ACCESS project the incidence of diabetes type 1 for databases with similar provenances was 13.1/100.000 PY. In primary care-only databases, the rates were similar as well: < 12.3 in BIFAP_PC, < 8.1 in CPRD, and 9.1/100.000 PY in ACCESS (6). The observed peak in CPRD in 2018 is probably due to a mix of prevalent and incident users at the start of the series. Overall, the IRs observed in SAFETY-VAC data sources are in the range of the meta-analysis of 193 studies results (15/100,000 PY) (75). The nature of PEDIANET database (specialist diagnoses and hospital admission in paediatric population) explains the higher rates of diabetes type 1 in comparison to the other data sources. For instance, rates ranged from 14.3 to 38.9/100.000 PY, with a sharp peak of 73.3/100.000 PY (CI 95% 20.4-185.6) in 2020. The reference ACCESS project rate was 19.8/100.000 PY.

Table 22 Standardized incidence rates of diabetes type 1 per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	31.3				5.9	9.5	8.4	0.0
2018	14.3				5.0	9.0	7.4	72.4
2019	15.6	9.4	9.0	6.4	5.2	9.2	7.4	6.4
2020	73.3	12.3	11.1	7.4	5.7	10.6	8.0	6.9
2021	38.9	7.9	7.9	7.6	6.6	10.4	9.2	8.1
2022	26.6	10.4	3.0	8.4	6.2	11.6	8.3	7.7
2023					5.7			

Figure 27. Standardized incidence rates of diabetes type 1 per data source.



10.1.4.9 Disseminated intravascular coagulation (DIC)

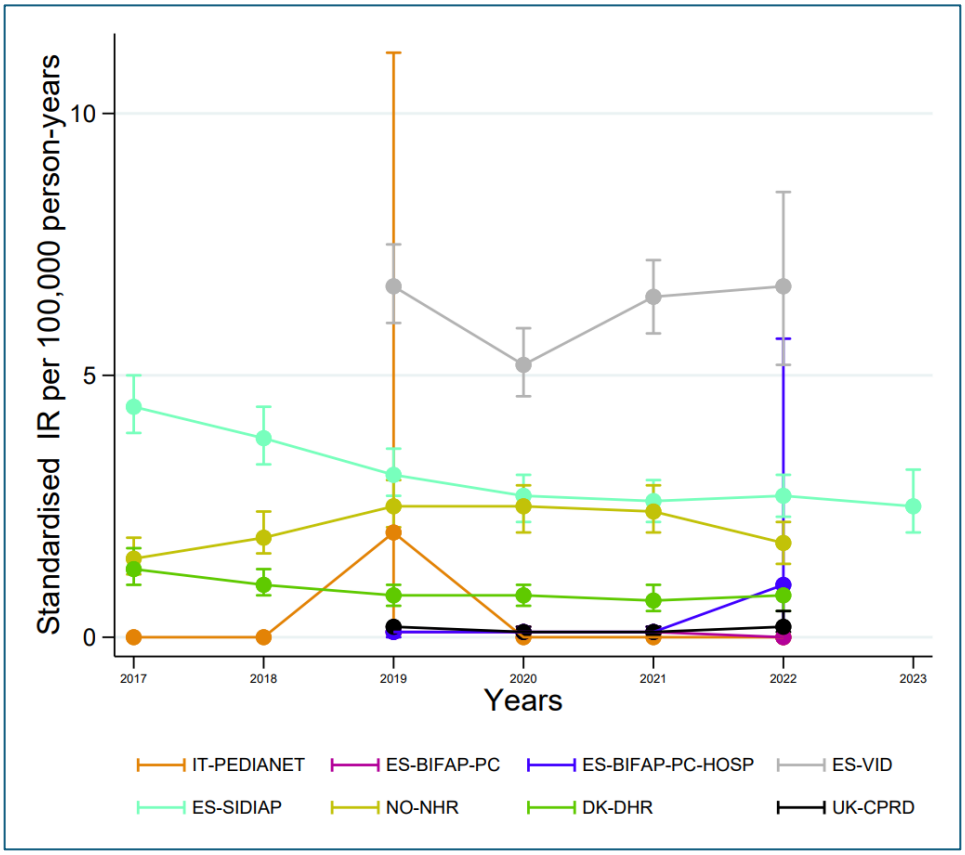
Disseminated intravascular coagulation (DIC) refers not to a single disease entity but a systemic pathophysiologic process. An overwhelming activation of the coagulation cascade triggers platelets and coagulation factors destruction causing microvascular fibrin thrombi. In turn, it can result in multiorgan dysfunction syndrome from tissue ischemia. Some conditions associated with acute DIC include septic shock, exsanguinating trauma, burns, or acute promyelocytic leukemia (79). The nature of this condition makes it a classic secondary hospital diagnosis.

The figure below shows differences among data sources with similar data provenances, i.e. primary care-hospital linkage. Although BIFAP_PC_HOSP includes hospital diagnoses, these are only primary diagnoses leading to incidence rates lower than in SIDIAP, VID, and NHR using primary and secondary diagnosis. SIDIAP and VID also included ICU and emergency room diagnosis. Overall, incidence rates are similar to CVM (22) and ACCESS rates. (6) Moreover, Cullen et al. (80) reported a pre-COVID-19 pandemic rate (2015-2019) of 1.0/100,000 PY, and a pandemic rate of 1.3/100,000 PY in Scottish population. These figures are similar to the rates in DHR, but higher to the rates in CPRD and BIFAP, and 1 to 5 times lower in comparison to VID, SIDIAP and NHR.

Table 23 Standardized incidence rates of DIC per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				4.4	1.5	1.3	0.0
2018	0.0				3.8	1.9	1.0	0.1
2019	2.0	0.1	0.1	6.7	3.1	2.5	0.8	0.2
2020	0.0	0.1	0.1	5.2	2.7	2.5	0.8	0.1
2021	0.0	0.1	0.1	6.5	2.6	2.4	0.7	0.1
2022	0.0	0.0	1.0	6.7	2.7	1.8	0.8	0.2
2023					2.5			

Figure 28 Standardized incidence rates of DIC per data source.



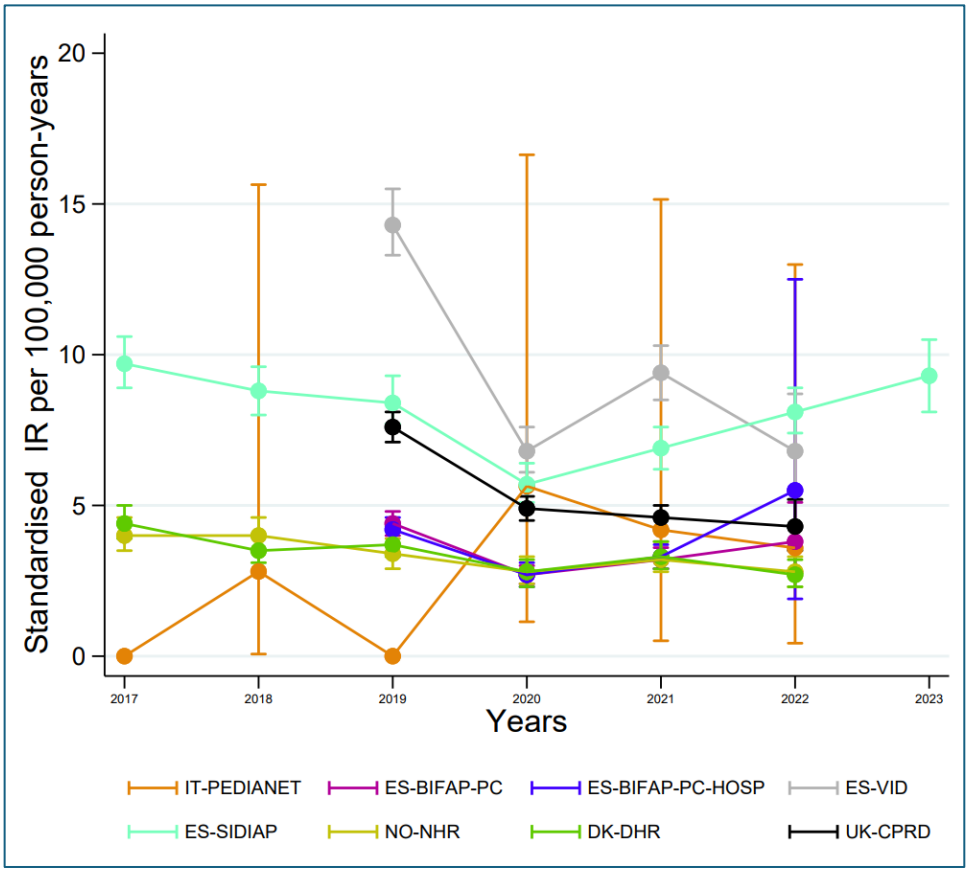
10.1.4.10 Erythema multiforme (EM)

Erythema multiforme (EM) is a skin and mucous membrane disease characterized by symmetric red, patchy lesions, primarily on the arms and legs (arms and forearms) and affecting mostly children and young adults. Of all cases of EM, approximately 20% occur in childhood (81). EM is more common in women compared with men (82). Although the cause is unknown, EM frequently occurs as an immunologic process initiated by a virus or medications such as anticonvulsants, sulfonamides, nonsteroidal anti-inflammatory drugs, and other antibiotics. EM is the mildest of related skin disorders: Stevens-Johnson syndrome and toxic epidermal necrolysis (TEN). A prior clinical definition form is available from ACCESS (83). EM is reported worldwide without any ethnic predilection. It occurs at any age, more frequently in young adults. Disease prevalence and annual incidence are reported to be unknown but rare in children (0.1 to 0.2/100000 PY for severe cases (84)), prevalence appears to be below 1% worldwide (82). Reported annual rates for EM requiring hospitalization varied from 0.1 to 4/100,000 PY across 9 different studies (82). In this study, EM has been mainly detected as primary, emergency and specialist diagnoses (**Annex 7**). Incidence rates are similar to 2019 CVM rates (22). Except in PEDIANET-IT (children), incidence rates dropped in 2020/2021 in all databases. It is probably due to the underutilization of services during the COVID-19 pandemic period. In general, EM rates are in line with the few data reported in the literature and, overall, hospital-linked data in addition to primary care data slightly increase the incidence of the detected cases.

Table 24 Standardized incidence rates of EM per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				9.7	4.0	4.4	0.0
2018	2.8				8.8	4.0	3.5	8.0
2019	0.0	4.4	4.2	14.3	8.4	3.4	3.7	7.6
2020	5.6	2.7	2.7	6.8	5.7	2.8	2.8	4.9
2021	4.2	3.2	3.3	9.4	6.9	3.2	3.3	4.6
2022	3.6	3.8	5.5	6.8	8.1	2.8	2.7	4.3
2023					9.3			

Figure 29. Standardized incidence rates of EM per data source.



10.1.4.11 Erythema nodosum (EN)

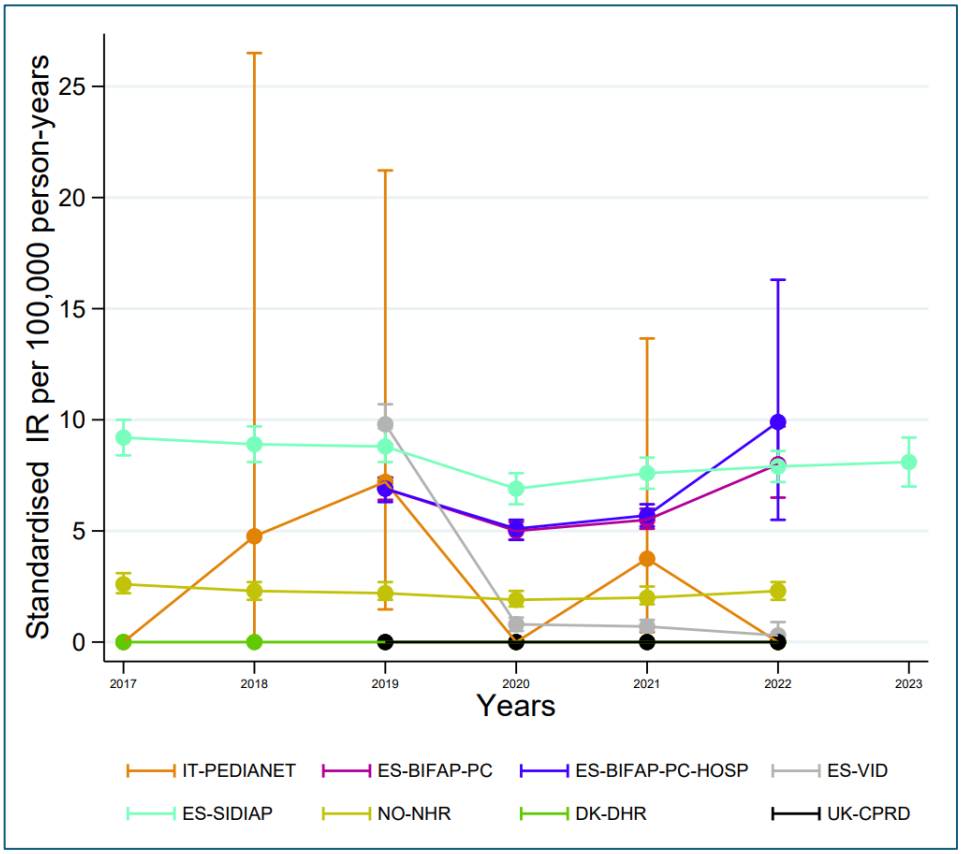
Erythema nodosum (EN) is a type of panniculitis, an inflammatory disorder affecting subcutaneous fat. It is a delayed-type hypersensitivity reaction that most often presents as erythematous tender nodules on the anterior shins. Less commonly, they affect the thighs and forearms. EN can occur secondary to a wide variety of conditions (85,86). EN occurs in approximately 1 to 5/100,000 persons, varying upon the different geographic areas and the various associated triggering diseases, and, in Europe, are estimated 14 /100,000 people per year (87–89).

Overall, we observed EN annual incidence rates that are in agreement with the reported worldwide and European estimates (<14/100000 PY). We observed values between 0.5 to 2.6/100,000 PY during the study period in NHR and between 5.0 to 9.9/100,000 PY during the study period in both BIFAP_PC and BIFAP_PC_HOSP. VID has values <1/100,000 PY in all years except 2019 (9.8). SIDIAP has generally slightly higher values during the whole study period (2017-2023) that range from 6.9 to 9.2/100,000 PY. PEDIANET shows annual IRs from null in 2017, 2020, 2022 to a peak of 7.2/100,000 PY in 2019. There were no EN cases detected in the current instance of DHR and CPRD, which may be due to the ETL process.

Table 25 Standardized incidence rates of EN per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				9.2	2.6	0.0	0.0
2018	4.8			0.0	8.9	2.3	0.0	0.0
2019	7.2	6.9	6.9	9.8	8.8	2.2	0.0	0.0
2020	0.0	5.0	5.1	0.8	6.9	1.9	0.0	0.0
2021	3.8	5.5	5.7	0.7	7.6	2.0	0.0	0.0
2022	0.0	8.0	9.9	0.3	7.9	2.3	0.0	0.0
2023					8.1			

Figure 30 Standardized incidence rates of EN per data source.



10.1.4.12 Generalized convulsion

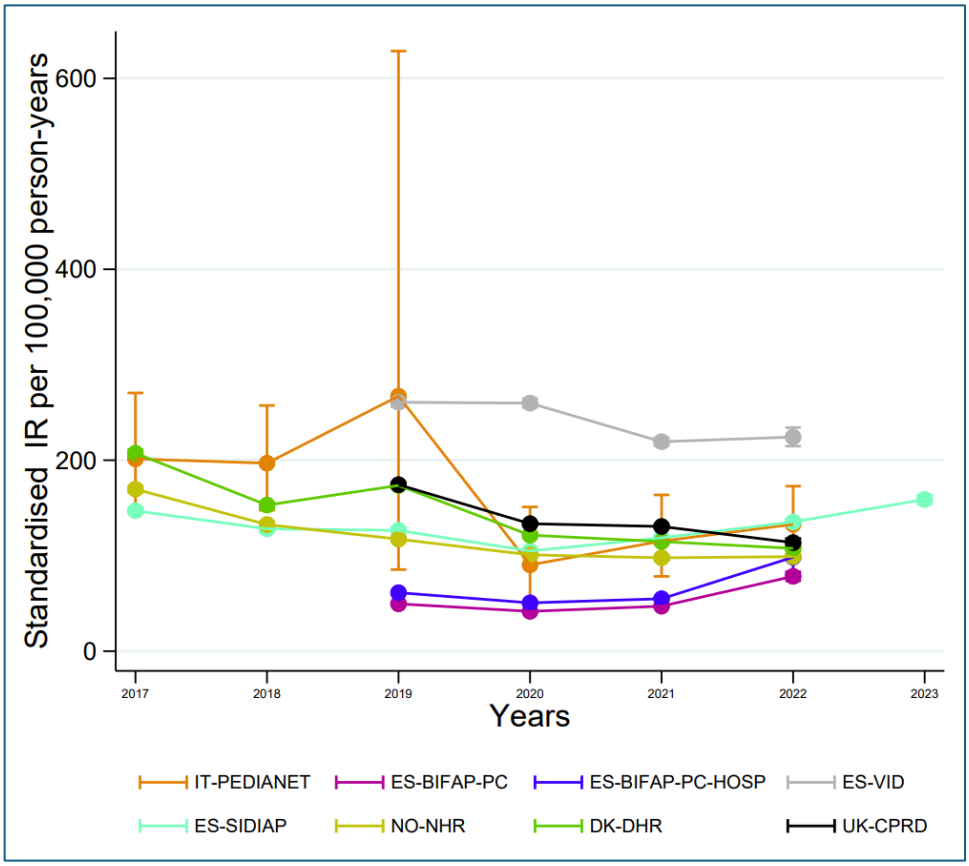
Seizures are paroxysmal alterations of neurologic functions caused by the excessive, hypersynchronous discharge of neurons in the brain, causing generalized convulsion. It usually results in sudden, involuntary muscular contractions, sensory disturbances, autonomic dysfunction and behavioural abnormalities, and impairment or loss of consciousness.

In a systematic review and meta-analysis of incidence studies in 2017, it has been reported a pooled incidence rate of epilepsy of 61.4/100000 PY (90). The incidence strongly varied from low/middle-income (139/100000 PY) to high-income countries (49/100000 PY). In ACCESS (6), the pooled incidence rate of generalized convulsion for primary care-only databases was 73.64/100.000 PY (CI 95% 43.77-103.51). This rate is similar to the rates reported in this study for BIFAP_PC: 49.6-78.4/100000 PY during the study period. In CPRD (primary care-only), rates are higher: 114 in 2022 to 278/100.000 PY in 2018 to ACCESS and to previous epidemiologic studies in the UK (91), but in line with general convulsions rates reported in 2015-2020 (245-278/100000 PY) (70). DHR rates ranges from 107.7 to 207.4/100000 PY during the study period, and these values are in line with previously reported incidence of generalized convulsions in Denmark (148/100000 PY) (70,92). The ACCESS incidence rate for generalized convulsion in databases combining primary care and hospital data was 152.35/100.000 (CI 95% 78.67-226.02). This rate is comparable to the incidences produced in this study, with exception of the slightly lower rates observed in BIFAP_PC_HOSP and in PEDIANET in 2019.

Table 26 Standardized incidence rates of generalized convulsion per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	201.4				147.1	169.7	207.4	0.0
2018	196.8				128.1	132.6	153.0	278.2
2019	267.1	49.6	61.3	260.7	126.5	117.3	173.6	174.4
2020	90.4	41.7	50.6	259.8	104.9	100.9	121.5	133.5
2021	115.3	47.1	54.9	219.4	118.7	97.8	114.8	130.6
2022	132.9	78.4	98.4	224.3	135.4	99.0	107.7	113.7
2023					158.9			

Figure 31 Standardized incidence rates of generalized convulsion per data source.



10.1.4.13 Graves' disease

Grave´s disease (GD) is a common form of hyperthyroidism with a diffuse hyperplastic goiter. It is an autoimmune disorder that produces thyrotropin-related antibodies (TRAb) against the thyroid stimulating hormone receptor. These autoantibodies activate the TSH receptor, thereby stimulating the thyroid gland and hypersecretion of thyroid hormones. These autoantibodies can also affect the eyes (Grave´s ophthalmopathy) and the skin (Grave´s dermopathy) (93).

Published literature on the incidence of Grave´s disease reports rates between 20 to 30 cases per 100,000 persons (94), it can go as low as 8.3/100,000 PY in Danish population in 2018 (95). Incidence rates in this study resulted higher than the mentioned values. In BIFAP-ES, Grave's disease rates were about 50/100,000 PY. In Norwegian population (NHR), disease rates reached an average rate of 67/100,000 PY. The highest incidence rates were reached in the SIDIAP and VID. The first accounted for rates above 100 till 123/100,000 PY, with a drop in 2020 (80/100,000 PY). In VID, Grave's disease rates increased from 104/100,000 PY in 2019 to 263 in 2020, 194 in 2021 and 158/100,000 PY in 2022. In general, there is a rise in the incidence in the years 2020 and 2022, this pattern has been also described in the CVM project for the event autoimmune thyroiditis (22).

In VID and SIDIAP, the diagnosis code that identified most of the cases was ICD10 E05.90 (thyrotoxicosis, unspecified without thyroid crisis or storm), which might be too broad as disease identifier, this must be further explored as an explanation to the higher rates in this study. Grave's disease codes were mainly identified as primary care diagnosis in BIFAP, VID, SIDIAP, and PEDIANET. In SIDIAP, hospitalization codes were also important. In NHR, specialized outpatient contact was the main source of diagnoses.

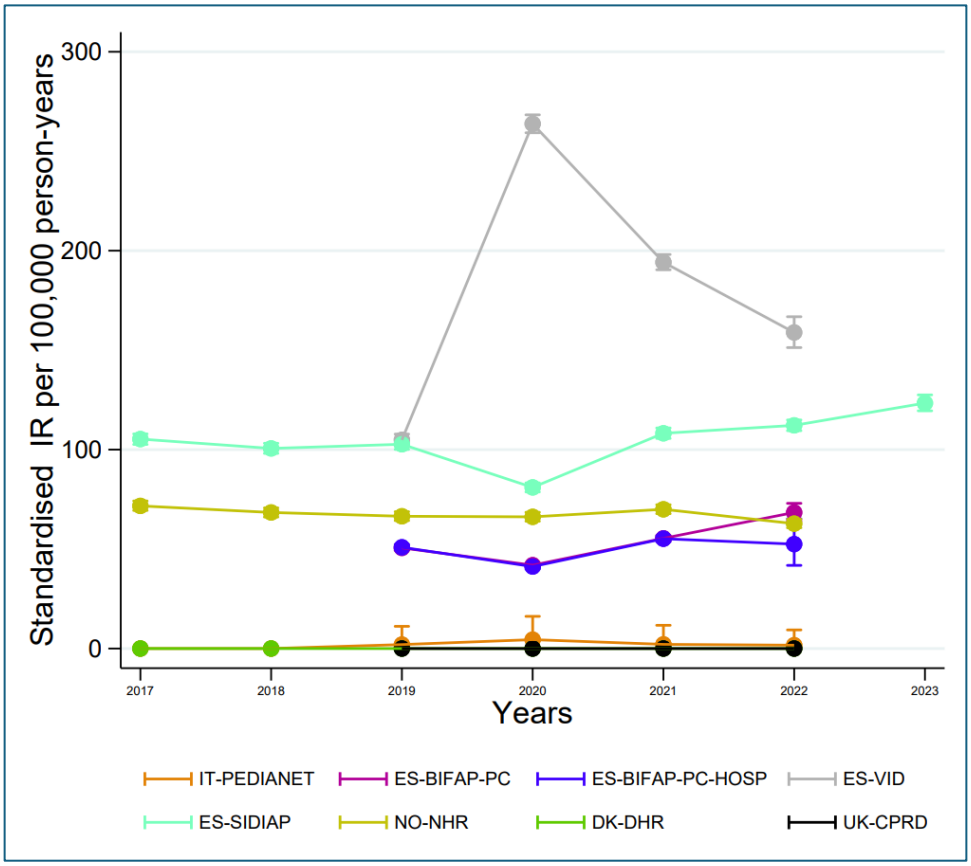
Grave's disease is a rare disease in children accounting for an incidence of 0.1/100000 PY in young children and 3/100,000 in adolescents from the UK and Ireland (96). In our study, the average rate in PEDIANET was 2/100,000 PY, with a peak up to 4.5 in 2020.

DHR and CPRD did not identified cases in the data instance used for this study.

Table 27 Standardized incidence rates of GD convulsion per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				105.3	71.7	0.0	0.0
2018	0.0				100.6	68.4	0.0	0.0
2019	2.0	50.6	50.9	104.9	102.7	66.5	0.0	0.0
2020	4.5	41.9	41.2	263.8	81.0	66.2	0.0	0.0
2021	2.1	55.4	55.2	194.2	108.2	70.0	0.0	0.0
2022	1.7	68.4	52.5	158.9	112.2	62.8	0.0	0.0
2023					123.4			

Figure 32 Standardized incidence rates of GD convulsion per data source



10.1.4.14 Guillain Barré Syndrome (GBS)

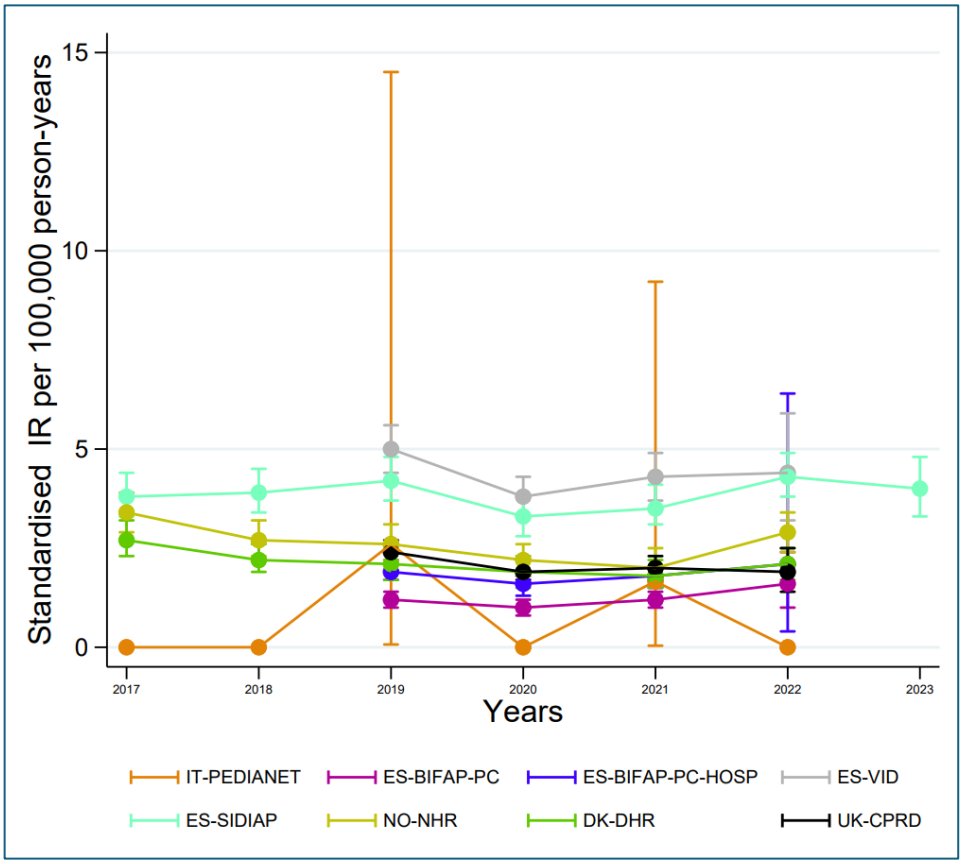
GBS is thought to be an immune-mediated disorder and a common cause of acute flaccid paralyses. Its aetiology and pathophysiology are not fully understood. Autoimmune antibodies and/or inflammatory cells cross-react targeting peripheral nerves and roots, resulting in their demyelination and/or axonal damage. This leads to sensory abnormalities, weakness in limbs or cranial nerve-innervated muscles, hypo- or areflexia, autonomic dysfunctions, and a cytoalbuminologic dissociation in the cerebrospinal fluid. A prior clinical definition form is available from ACCESS (97).

It is reported that GBS occurs worldwide with an overall incidence rate of around 1-2/100,000 PY, affecting all age groups, but is slightly more common in males than in females (98). Mortality, or severe disability due to GBS, occurs in around 20% of GBS affected people. Incidence rates reported in this study vary by type of data source, since it is mainly diagnosed in hospital, but are similar to the rates obtained in the previous CVM (22) and ACCESS (6) studies, as well as to worldwide overall rates published elsewhere (70,98–105).

Table 28 Standardized incidence rates of GBS per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0			0.0	3.8	3.4	2.7	0.0
2018	0.0			0.0	3.9	2.7	2.2	3.2
2019	2.6	1.2	1.9	5.0	4.2	2.6	2.1	2.4
2020	0.0	1.0	1.6	3.8	3.3	2.2	1.9	1.9
2021	1.7	1.2	1.8	4.3	3.5	2.0	1.8	2.0
2022	0.0	1.6	2.1	4.4	4.3	2.9	2.1	1.9
2023					4.0			

Figure 33 Standardized incidence rates of GBS per data source.



10.1.4.15 Haemorrhagic stroke

Haemorrhagic stroke is defined as an abrupt rupture of a weakened blood vessel in the brain causing a sudden cognitive loss of neurological function. Haemorrhagic strokes include intracerebral haemorrhage (ICH, bleeding within the brain) and subarachnoid haemorrhage (SAH, bleeding between the inner and outer layers of tissue covering the brain within the subarachnoid space). ICH and SAH account for approximately 10% and 3% of all strokes, respectively. Estimates of ICH incidence around the world varies but have generally ranged from 10 to 20 per 100,000 PY, whereas SAH incidence was reported to be 9.7 per 100,000 PY (106).

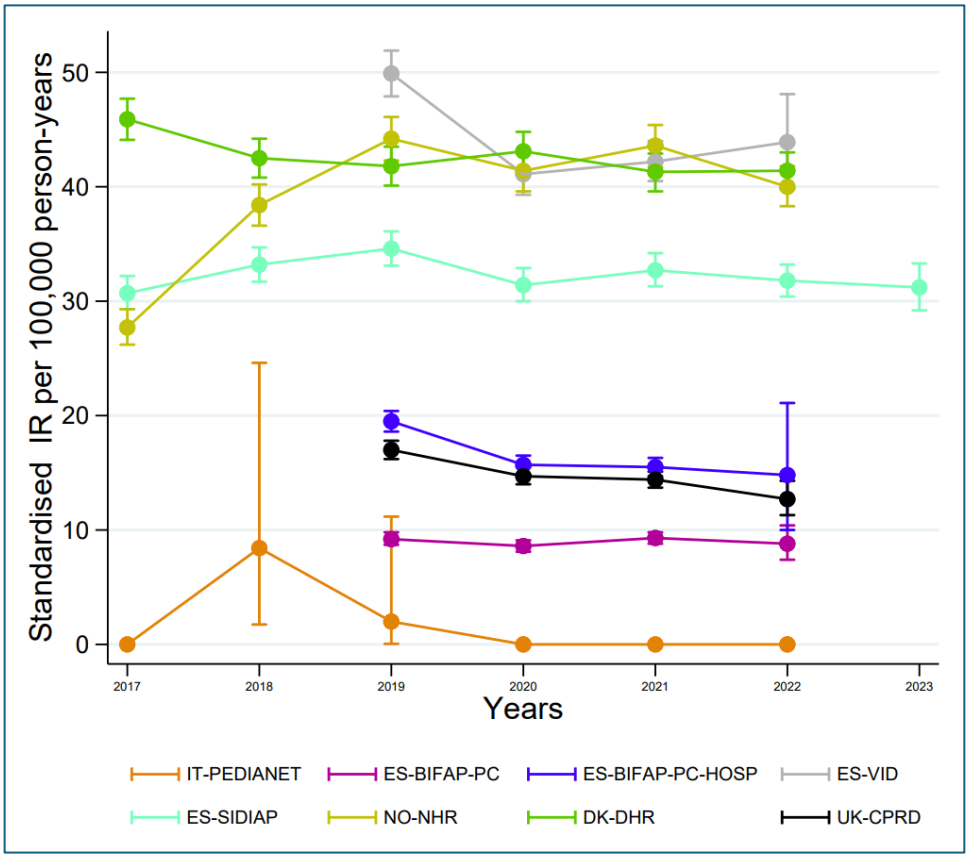
Different studies provided information on the overall annual incidence of haemorrhagic strokes applying heterogeneous methodologies that may contribute to the differences from one territory to another. Across different countries, the reported median crude incidence is 29/100000 PY. In European countries, the haemorrhagic strokes incidence per 100000 PY varied from 9-12 in Spain, 12 in Germany, 15 in Italy, 26 in England and 43 in France (107).

According to the code counts (**Annex 7**), cases of haemorrhagic stroke are detected in primary and hospital care settings in our data sources. Overall, the SAFETY-VAC rates are in line with the ones previously reported in Europe, ranging from 9 to 43/100000 PY, as listed above. In the reference incidence rate produced in ACCESS project (6), rates in the combined primary care-hospital databases were 43.58/100.000 PY, which is similar to the rates in VID, SIDIAP, DHR and NHR. Rates in BIFAP_PC_HOSP are lower: <20/100.000 PY. In primary care-only databases, both BIFAP_PC and CPRD reported similar rates to the corresponding ACCESS rate 10.54/100.000 PY.

Table 29 Standardized incidence rates of haemorrhagic stroke per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				30.7	27.7	45.9	0.0
2018	8.4				33.2	38.4	42.5	18.4
2019	2.0	9.2	19.5	49.9	34.6	44.2	41.8	17.0
2020	0.0	8.6	15.7	41.1	31.4	41.4	43.1	14.7
2021	0.0	9.3	15.5	42.2	32.7	43.6	41.3	14.4
2022	0.0	8.8	14.8	43.9	31.8	40.0	41.4	12.7
2023					31.2			

Figure 34. Standardized incidence rates of haemorrhagic stroke per data source.



10.1.4.16 Hashimoto's thyroiditis

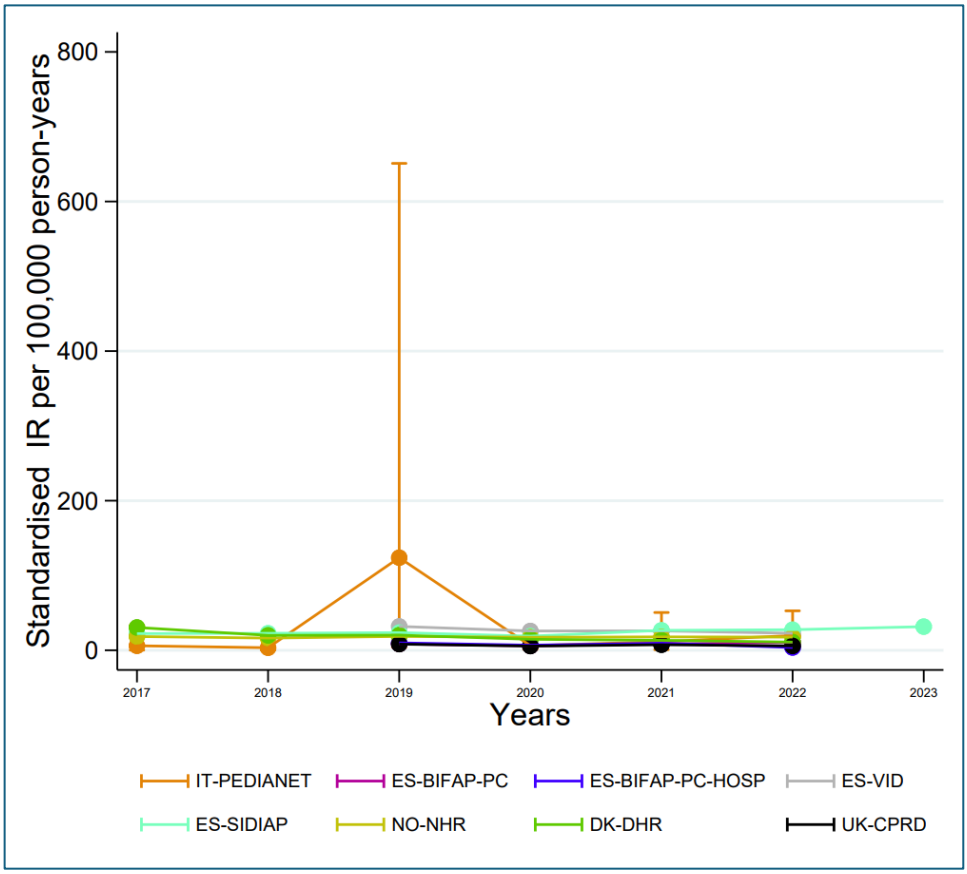
Hashimoto thyroiditis (HT), also known as chronic autoimmune thyroiditis or chronic lymphocytic thyroiditis, is a chronic inflammation of the thyroid gland. It is considered the most common autoimmune and endocrine disorder, as well as the most common cause of hypothyroidism. Based on aetiology, HT can be classified into primary and secondary forms. Most HT ultimately evolve into hypothyroidism, although at diagnosis time patients can be euthyroid or even hyperthyroid (108).

Worldwide, several studies to analyze the IRs of HT have been carried out based on different designs and laboratory analyses. Results highly vary among areas, age and sex. In European countries, IR of autoimmune hypothyroidism ranges from 47.1/100,000 PY in women and 2.1/100,000 in men in Vigo (Northwest Spain) between 1990 and 1992 (109) to 448/100,000 PY for women and 92/100,000 PY for men within the population of Tayside, Scotland (approximately 390,000 people) over eight years (110). In Denmark, the age-standardized pre-iodization rate of spontaneous hypothyroidism was 44.4/100,000 PY in women and 11.9/100,000 PY in men, rising markedly with age. The post-iodization incidence of overt hypothyroidism increased by a statistically significant 20 % in females and 40 % in males (111). Another Danish study showed an IR of 32.8/100,000 PY (112). Overall, the SAFETY-VAC rates are lower than the ones reported in Europe. VID and SIDIAP accounted for rates lees than 31/100,000 PY during the study period, both are linked primary care-hospital databases, with hospital (primary and secondary diagnosis) being an important provenance of diagnosis. In fact, in SIDIAP the hospital secondary diagnoses contribute 33% to the total cases identified, see code counts in **Annex 7**. The percentage of hospital-related codes in BIFAP_PC-HOSP is 8%, but only primary diagnosis, which could explain the lower rates in this database (<10/100,000 PY). A similar IR of less than 10/100,000 PY is observed in primary care-based databases (BIFAP_PC and CPRD). In NHR and DHR, the average IR is 17/100,000 PY, with exception of 2017 in DHR. In paediatric population (PEDIANET), rates vary between 3.4 to 19.7/100,000 PY, with an important peak in 2019.

Table 30 Standardized incidence rates of Hashimoto's thyroiditis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	6.0				22.3	18.4	30.5	0.0
2018	3.4				22.7	16.3	19.9	10.6
2019	123.8	8.4	9.9	31.9	23.6	18.5	20.2	8.2
2020	6.2	5.9	6.8	25.8	18.9	17.4	14.7	5.6
2021	11.4	8.4	9.5	25.9	26.6	17.9	13.3	7.4
2022	19.7	10.6	3.7	22.7	27.5	17.6	10.9	5.8
2023					31.6			

Figure 35 Standardized incidence rates of Hashimoto's thyroiditis per data source.



10.1.4.17 Idiopathic thrombocytopenic purpura (ITP)

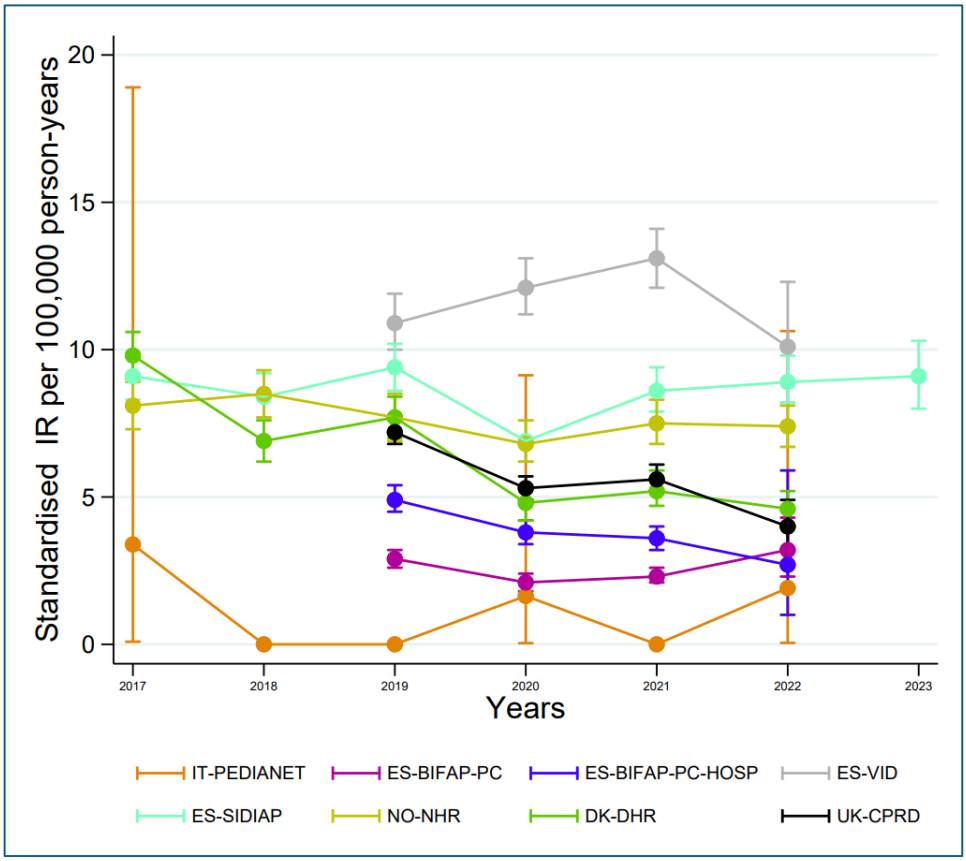
Thrombocytopenia is an abnormally low platelet count. Pathogenic mechanisms include insufficient production, abnormal distribution, or excessive destruction of platelets. Excessive destruction can be caused by microangiopathy, hereditary platelet abnormalities, or immunologic mechanisms. Idiopathic TP (ITP) refers to TP without an identified etiology, although an autoimmune etiology is frequently suspected but not always verified through exhaustive exclusion of differential diagnoses. A prior clinical definition form which describes thrombocytopenia and ITP is available from ACCESS (113).

In 2021, Willame C., et al., published an ITP incidence rate of 3.8/100,000 PY using a narrow selection of diagnosis codes in databases from Denmark, Spain, Italy, and the UK (50). This result is comparable to the rates in BIFAP and PEDIANET. It is, however, lower than the results obtained in the remaining data sources. The Nordic countries included in this study, Norway and Denmark, accounted for similar ITP IRs: 4.6 to 9.8/100,000 PY across the study period. These rates are higher in comparison to the ones reported for chronic ITP in 3 Nordic countries (1.8 to 2.8/100,000 PY) (114), but similar to the ITP rates in Canada (115) and in the US (116). Finally, the highest rates are observed in VID and SIDIAP (between 8.4 to 13.1/100,000 PY), which are comparable to the rates reported by Li et al. in children and adult population (68). In general, the ITP incidence rates in this study are stable in all data sources and across time. (117)

Table 31 Standardized incidence rates of idiopathic thrombocytopenic purpura per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	3.4				9.1	8.1	9.8	0.0
2018	0.0				8.4	8.5	6.9	10.7
2019	0.0	2.9	4.9	10.9	9.4	7.7	7.7	7.2
2020	1.6	2.1	3.8	12.1	6.9	6.8	4.8	5.3
2021	0.0	2.3	3.6	13.1	8.6	7.5	5.2	5.6
2022	1.9	3.2	2.7	10.1	8.9	7.4	4.6	4.0
2023					9.1			

Figure 36. Standardized incidence rates of idiopathic thrombocytopenic purpura per data source.



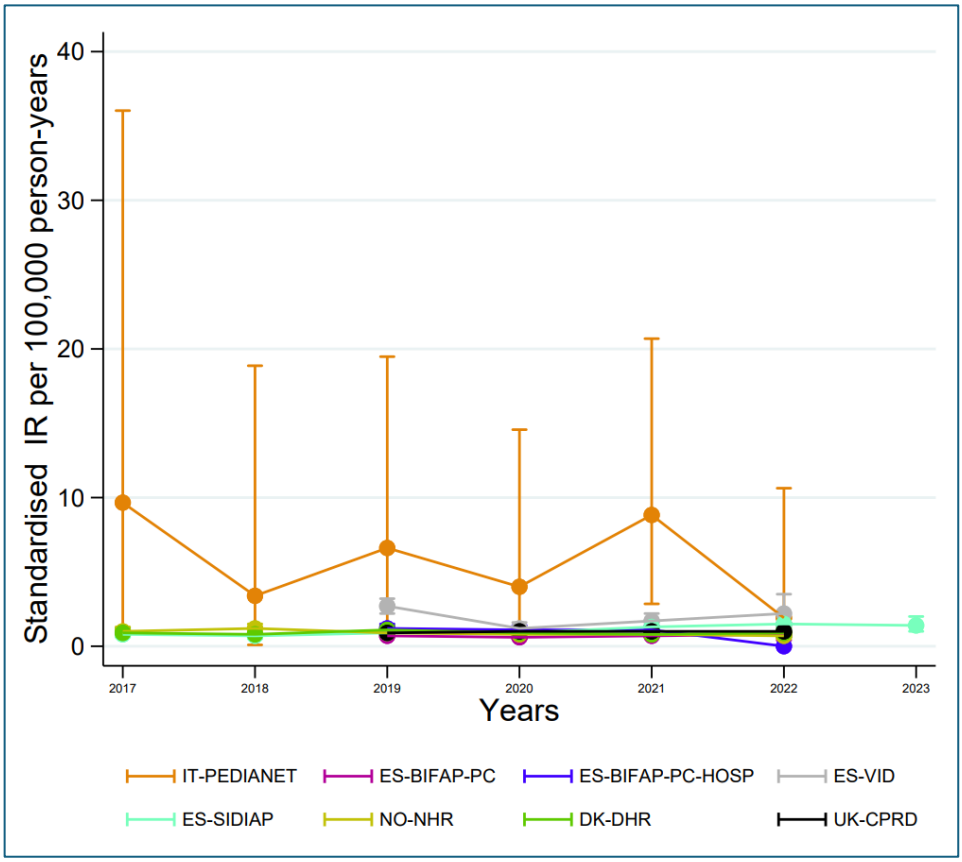
10.1.4.18 Kawasaki disease (KD)

Kawasaki disease (KD) is a systemic vasculitis of infancy and childhood affecting medium-sized muscular arteries. It causes potentially life-threatening changes in the coronary arteries of some children and is the most common cause of paediatric-acquired heart disease in high income countries. It primarily affects children younger than 5 years of age (80% of cases), although older children and teenagers can also get KD rarely. The syndrome is more common in boys than girls. Although KD's aetiology is unknown, it is suspected to be caused by a virus (117). The annual incidence of KD in Europe is about 10-15 per 100,000 children under 5 years (118). This is consistent with the incidence rates of KD in 0-1 and 2-4 years of age in this study (see **Annex 5** for age specific rates). Across the overall population, the standardized incidence rates are lower, except for PEDIANET, which includes only children.

Table 32 Standardized incidence rates of Kawasaki's disease per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	9.7				0.8	1.0	0.9	0.0
2018	3.4				0.7	1.2	0.8	1.1
2019	6.6	0.7	1.2	2.7	0.9	0.9	1.1	0.9
2020	4.0	0.6	1.1	1.2	0.9	0.8	0.9	1.0
2021	8.8	0.7	1.1	1.7	1.3	0.8	0.8	1.0
2022	1.9	0.8	0.0	2.2	1.5	0.7	0.9	1.0
2023					1.4			

Figure 37. Standardized incidence rates of Kawasaki's disease per data source.



10.1.4.19 (Meningo)encephalitis

Encephalitis defines an inflammation of the parenchyma of the brain. The presence of inflammation, oedema, and neuronophagia (neuronal cell death) is identified by histopathology in the pathologic diagnosis. Meningoencephalitis is diagnosticated with the focal accumulations of a mixed inflammatory cell infiltrate in the meninges and brain and is characterized by necrosis of brain parenchyma (with all cellular elements affected, especially in the periventricular region, and often associated with calcification), reactive microglial and astroglial proliferation, and the occurrence of enlarged cells (neuronal and glial elements) with intranuclear inclusions (119).

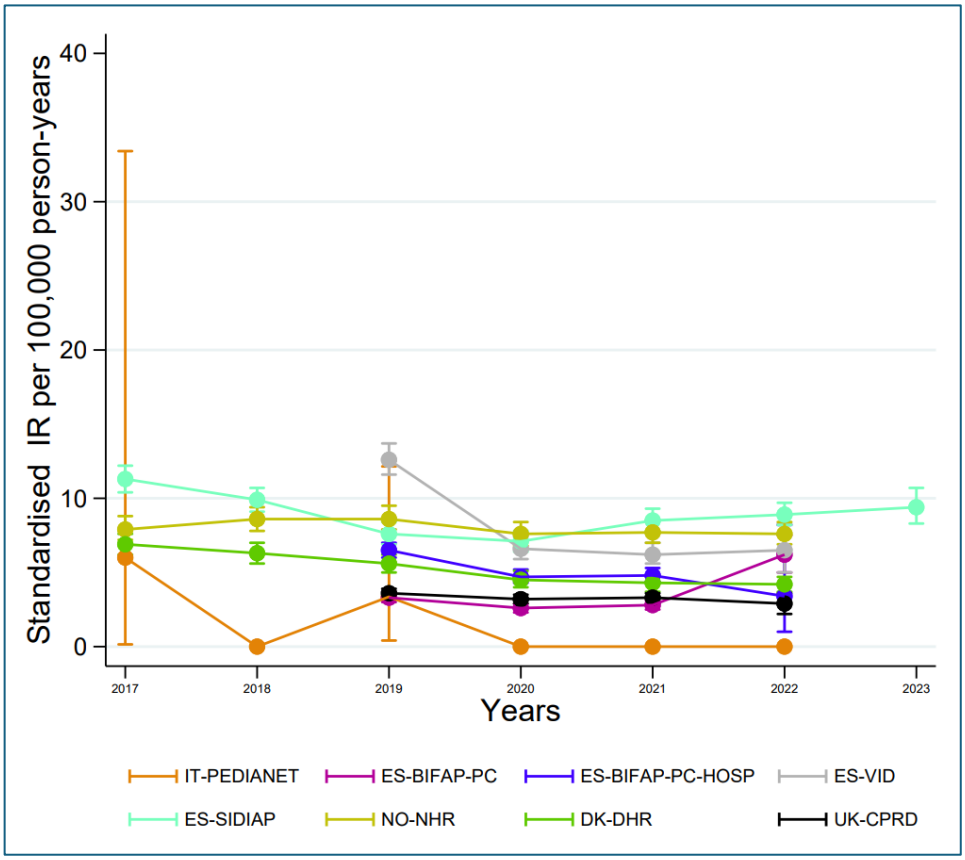
Overall, our results here show standardized annual IRs (100,000 PY) comparable with ACCESS results and Gubernot et al. (116) rates of 6.9-7.3/100,000 in the general US population and the reported incidence of 7.3 encephalitis hospitalizations per 100,000 population in the U.S. from 2000–2010 (120).

This event is mainly detected through hospitalization, emergency rooms and ICU settings, particularly in case of severe episodes. In fact, BIFAP_PC shows rates from 2.6 to 3.3 between 2019-2021, similarly to CVM and ACCESS rates for primary care-only data provenance, but also a peak of 6.2 in 2022. CPRD-UK (mainly primary care-only data) rates also go from 2.9 to 3.9 between 2018-2022 and are in line with both CVM and ACCESS rates. Considering data including hospitalization information, BIFAP_PC_HOSP ranges from 3.4 to 7.1 between 2018-2022, in line with CVM and ACCESS values. VID IRs are in line with CVM values for 2019 (12.6), and lower rates (6.2-6.6) are observed between 2020-2022 that are comparable with ACCESS values. SIDIAP rates range from 7.1 to 11.3 between 2017-2023 and are comparable with CVM and ACCESS rates. NHR rates were very high in CVM (22.9) but in this report are stable across 2017-2022 (7.6 to 8.6) and in line with ACCESS rates. PEDIANET-IT reported IRs for 2017 (6.0) and 2019 (3.4), which are comparable to ACCESS rates, but no cases are reported for the other years of the study period. DHR reported slightly lower rates than the other data sources with hospitalization data, ranging from 2.9 to 3.9/100000 PY, but overall, in line with other published studies.

Table 33 Standardized incidence rates of meningoencephalitis per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	6.0				11.3	7.9	6.9	0.0
2018	0.0				9.9	8.6	6.3	3.9
2019	3.4	3.3	6.5	12.6	7.6	8.6	5.6	3.6
2020	0.0	2.6	4.7	6.6	7.1	7.6	4.5	3.2
2021	0.0	2.8	4.8	6.2	8.5	7.7	4.3	3.3
2022	0.0	6.2	3.4	6.5	8.9	7.6	4.2	2.9
2023					9.4			

Figure 38. Standardized incidence rates of meningoencephalitis per data source.



10.1.4.20 Microangiopathy

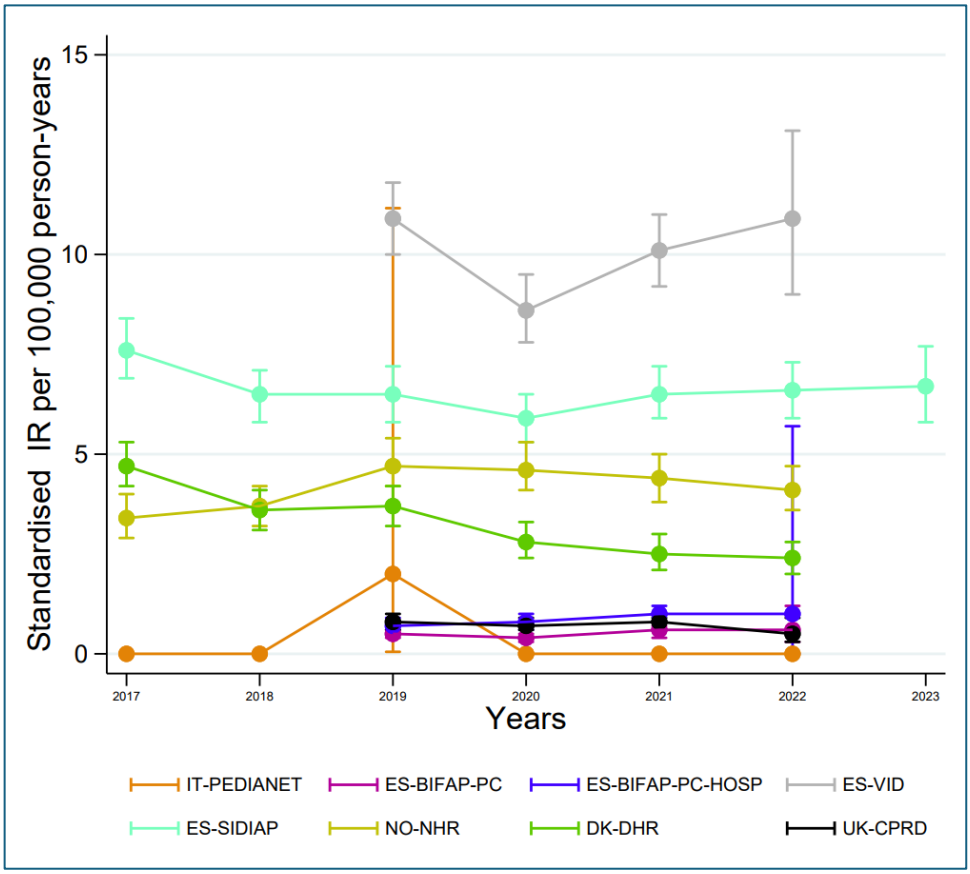
Microangiopathy is a disease that affects small blood vessels in the microcirculation, leading to microvascular dysfunction which can trigger different clinical scenarios. Cardiac microangiopathy can manifest through events of ischemic heart disease in the absence of angiographically significant coronary atherosclerosis, causing inflammation and/or abnormal vasomotor regulation, or through inadequate post-percutaneous coronary intervention (PCI) and/or -thrombolysis coronary reperfusion, including micro-embolic mechanism, or in the context of epicardial vessel disease (121).

Several conditions are associated with thrombotic microangiopathy (TMA) and no overall references about incidence of these conditions as a disease are reported. In this study, BIFAP_PC_HOSP incidence rates were <1/100,000 PY along the study period (primary diagnosis only), which is aligned to the rates reported in the CVM project. In contrast to the low rates of BIFAP_PC_HOSP, data including hospital and/or emergency rooms information from NHR, DHR, VID and SIDIAP reported higher rates between 2.4 and 10.9/100,000 PY, similar to the 3.3/100,000 reported in the ACCESS project. Primary care-only databases (CPRD and BIFAP-PC) reported rates less than 1/100000 PY, also aligned to the 0.32/1000,000 PY in ACCESS. Finally, PENIANET reported an incidence of 2/100000 PY in 2019 only.

Table 34 Standardized incidence rates of microangiopathy per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				7.6	3.4	4.7	0.0
2018	0.0				6.5	3.7	3.6	1.0
2019	2.0	0.5	0.7	10.9	6.5	4.7	3.7	0.8
2020	0.0	0.4	0.8	8.6	5.9	4.6	2.8	0.7
2021	0.0	0.6	1.0	10.1	6.5	4.4	2.5	0.8
2022	0.0	0.6	1.0	10.9	6.6	4.1	2.4	0.5
2023					6.7			

Figure 39. Standardized incidence rates of microangiopathy per data source.



10.1.4.21 Multiple sclerosis

Multiple sclerosis is an inflammatory demyelinating disease affecting the central nervous system, thought to result from the interaction of genetic and environmental factors that remain only partially understood. Pathologic findings include multiple sharply demarcated areas of demyelination throughout the white matter of the central nervous system. Clinical manifestations include visual loss, extra-ocular movement disorders, paresthesia, loss of sensation, weakness, dysarthria, spasticity, ataxia, and bladder dysfunction (122,123).

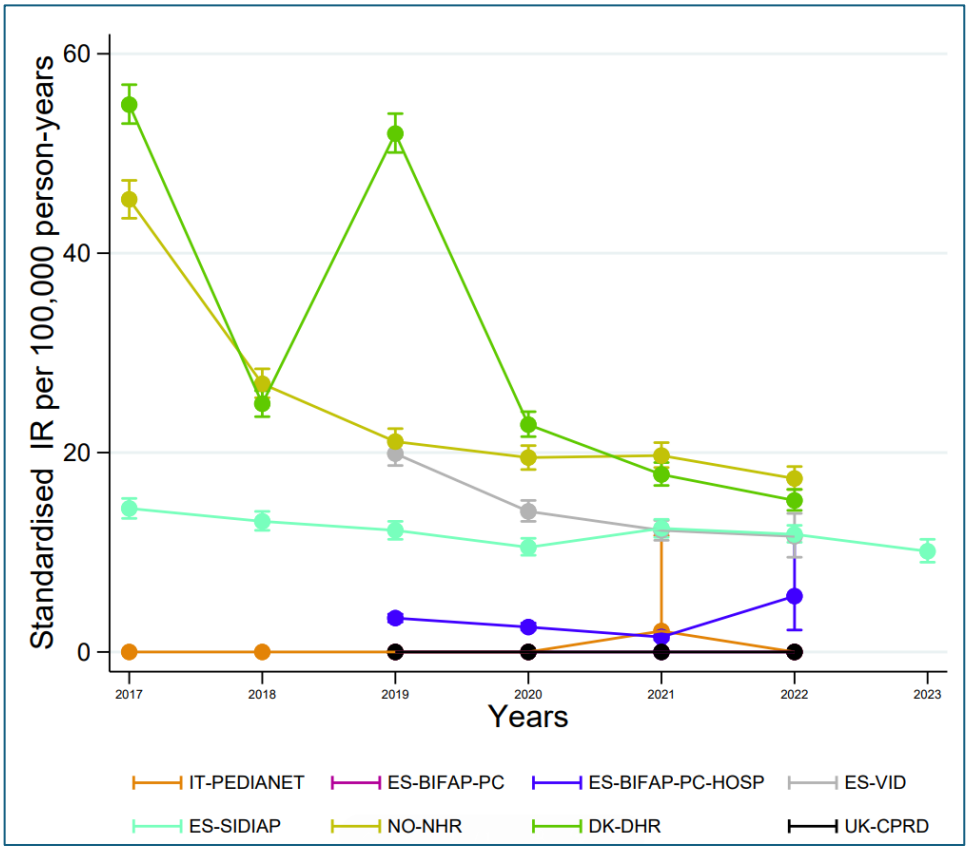
The estimated European mean annual multiple sclerosis IR ranged from 1.1 to 7.0 per 100,000 population (124,125), with variation across regions and countries. Published studies reported mean annual IRs per 100,000 population that varied from 8.3-11.5 in Denmark (126), 4.2-13.9 in Norway, 3.7-9.2 in Finland, 4.9-13.9 in the UK, to 5.3-6.5 in France, 2.6-17.4 in Italy, and 2.8-13.9 in Spain (127).

In our study, BIFAP_PC_HOSP and SIDIAP reported annual IRs in line with previously mentioned values, whereas VID had values in line with reported rates from 2020 to 2022, but higher values in 2019 (19.8/100,000 PY, respectively). NHR showed higher IRs than previously reported during the study period, ranging from 17.4 in 2022 to 45.4/100,000 PY in 2017. The same situation is true for DHR, with IRs from 15.2 to 54.9 during the study period. Almost null IRs are reported by BIFAP_PC, as well as for PEDIANET, which shows a peak of 2.1/100,000 PY only in 2021. CPRD does not have this event in the current data instance.

Table 35 Standardized incidence rates of multiple sclerosis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				14.4	45.4	54.9	
2018	0.0				13.1	26.9	24.9	
2019	0.0	0.0	3.4	19.9	12.2	21.1	52.0	
2020	0.0	0.0	2.5	14.1	10.5	19.5	22.8	
2021	2.1	0.0	1.5	12.2	12.4	19.7	17.8	
2022	0.0	0.0	5.6	11.6	11.8	17.4	15.2	
2023					10.1			

Figure 40 Standardized incidence rates of multiple sclerosis per data source.



10.1.4.22 Myocarditis

Myocarditis is an inflammatory disease of the myocardium that encompasses several different diseases. It is widely heterogeneous both regarding its aetiologies and its form of clinical presentation. It is frequently caused by viral and nonviral infections or post-viral immune-mediated responses or non-infectious triggers (autoimmune diseases, hypersensitivity reactions to drugs, toxic reactions to drugs, toxins, etc.). Diagnosis is established by histological, immunological, and immunohistochemical criteria, which may differ with respect to the appearance under the microscope and to clinical aetiology (128).

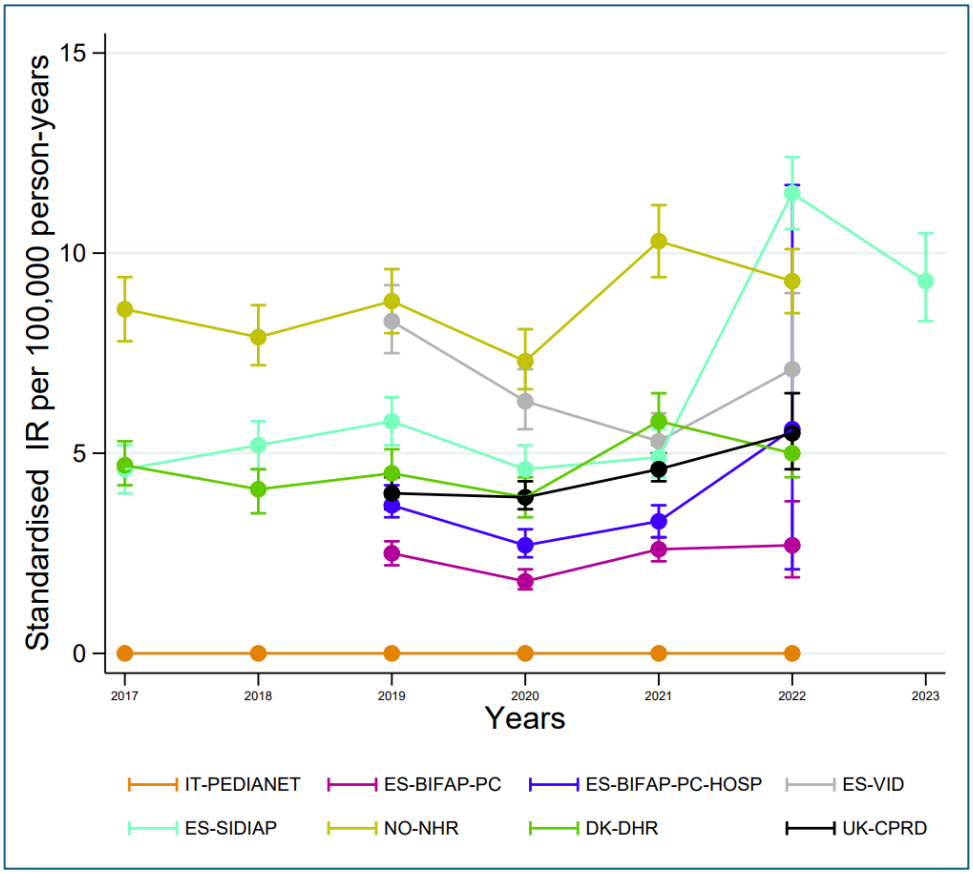
The true incidence of myocarditis is challenging to quantify due to the variability in clinical presentation (129). From the Global Burden of Disease (GBD) study, age-standardize IRs of myocarditis across 204 countries and territories are reported to vary from 16.7 in 1990 to 16.0/100000 PY in 2019, with European values ranging from 16.3 to 17.7/100000 PY (130). The incidence globally increases with age, from 5.9 in to 150.0/100000 PY in <5 to +95 years old individuals, respectively.

Across 4 Nordic countries (Denmark, Finland, Norway, and Sweden), the myocarditis IRs were reported to be 9.7 for males and 4.3/100,000 PY for females (131,132). The myocarditis background rates recently calculated during the Global COVID Vaccine Safety (GCoVS) project within ≈197 million people ranged from 1.6 to 7.7/100,000 PY (70). Overall, SAFETY-VAC data sources reported myocarditis IRs in line with these Nordic Countries' and GCoVS studies' estimates. Moreover, our study shows that most of the standardized annual IRs are comparable with the CVM study-reported values (22). All data sources reported rates (100,000 PY) below 10, with some exceptions in some years. BIFAP_PC rates are 1.8-2.7 between 2019-2022, in line with CVM values and primary care-only values reported in ACCESS (6). BIFAP_PC_HOSP rates (2.7-3.7) between 2019-2021 are in line with CVM rates, but with smaller higher rates shown in 2022 (5.6). CPRD rates (3.9-4.6) between 2018-2022 are in agreement with CVM rates. VID rates (5.3-8.6) are comparable with CVM results and stable across the study period 2018-2022. SIDIAP rates (4.6-5.8) in 2018-2021 are also comparable with CVM results, except for peaks of 11.5 and 9.3 in 2022 and 2023 respectively. NHR rates range from 7.3 to 10.3 between 2017-2022, and are slightly higher than background rates of the CVM study (7.0). DHR shows stable rates ranging from 3.9 to 5.8 across the study periods. Both NHR and DHR rates are in line with Nordic countries published estimates. PEDIANET does not have cases.

Table 36 Standardized incidence rates of myocarditis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				4.6	8.6	4.7	0.0
2018	0.0				5.2	7.9	4.1	4.4
2019	0.0	2.5	3.7	8.3	5.8	8.8	4.5	4.0
2020	0.0	1.8	2.7	6.3	4.6	7.3	3.9	3.9
2021	0.0	2.6	3.3	5.3	4.9	10.3	5.8	4.6
2022	0.0	2.7	5.6	7.1	11.5	9.3	5.0	5.5
2023					9.3			

Figure 41. Standardized incidence rates of myocarditis per data source.



10.1.4.23 Narcolepsy

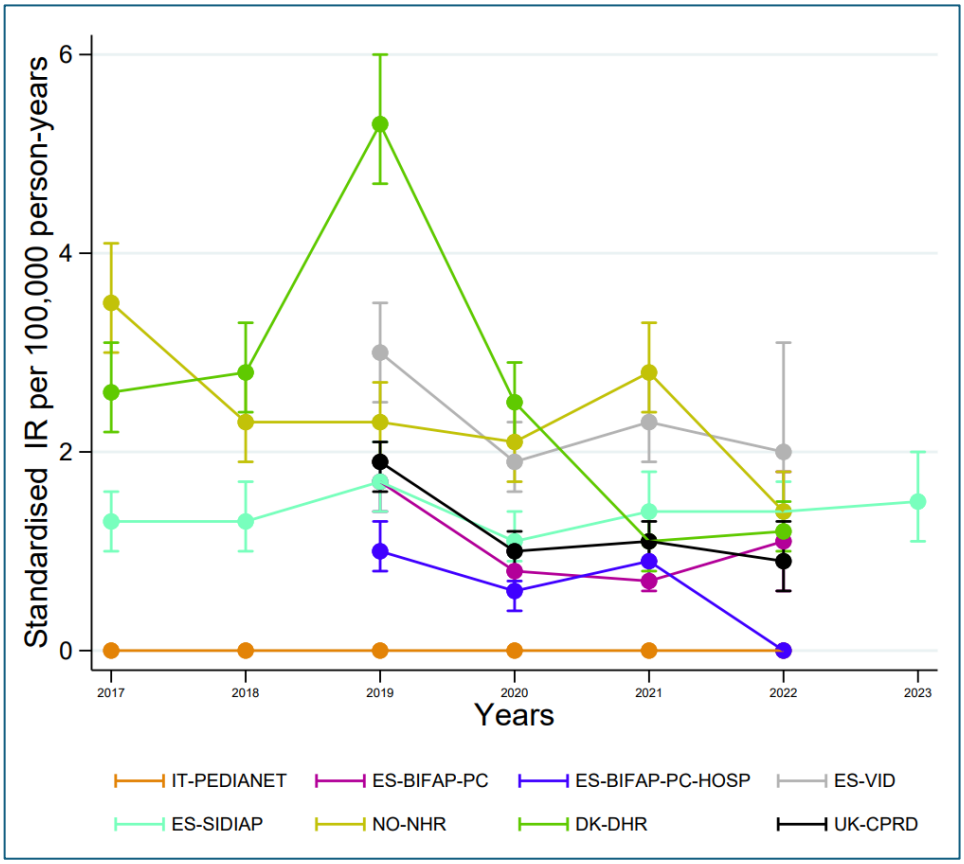
Narcolepsy is a chronic neurological disorder where the brain is not able to control sleep-wake cycles. Individuals with narcolepsy often have excessive daytime sleepiness and cataplexy episodes of muscle weakness brought on by emotions. They may also experience uneven/interrupted sleep that can involve waking up frequently during the night and frequently entering REM sleep rapidly, within 15 minutes of falling asleep. Other symptoms may include hypnagogic hallucinations, sleep paralysis, fragmented nocturnal sleep, as well as impaired ability for sustained attention, and non-sleep symptoms such as obesity, anxiety, cognitive and emotional disturbances, behavioural problems, and early puberty in children (133).

Narcolepsy standardized annual IRs in this study are in line with literature reported values (1-2/100,000 PY) ACCESS project (6), and CVM study results (22). Rates (100,000 PY) are 0.8 to 1.7 in BIFAP_PC between 2019-2022, 0 to 1.0 in BIFAP_PC_HOSP between 2019-2022, 0 to 3.0 in VID between 2017-2022, 1.1 to 1.7 in SIDIAP between 2017-2023, 1.4 to 3.5 in NHR between 2017-2022, 1.1 to 5.3 in DHR between 2017 and 2021, and 0 to 3.0 in CPRD between 2017-2022. No cases were observed in PEDIANET (paediatric cohort). In general, these figures are similar to the incidence rates reported for US (1.37/100,000 persons) and (134) Scottish population (0.6/100,000 population) (135), and to the rates published by Dodd C., et al., (136) for several Canadian, European and Taiwanese populations (0.22 to 1.52 per 100,000 person-years).

Table 37 Standardized incidence rates of narcolepsy per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				1.3	3.5	2.6	0.0
2018	0.0				1.3	2.3	2.8	3.0
2019	0.0	1.7	1.0	3.0	1.7	2.3	5.3	1.9
2020	0.0	0.8	0.6	1.9	1.1	2.1	2.5	1.0
2021	0.0	0.7	0.9	2.3	1.4	2.8	1.1	1.1
2022	0.0	1.1	0.0	2.0	1.4	1.4	1.2	0.9
2023					1.5			

Figure 42. Standardized incidence rates of narcolepsy per data source.



10.1.4.24 Pancreatitis, acute

Pancreatitis is an inflammatory condition where digestive enzymes damage the pancreas, causing either an acute or chronic condition. Common causes of acute pancreatitis are gallstones, heavy alcohol abuse, direct trauma, certain medications, infections, or tumours. The acute form may evolve into chronic due to heavy alcohol consumption, high levels of blood fats, high blood calcium, or certain genetic disorders, such as cystic fibrosis. Symptoms of pancreatitis include pain in the upper abdomen, nausea, and vomiting. The diagnosis of acute pancreatitis requires abdominal pain, three times greater serum lipase activity (or amylase activity), and characteristic medical imaging findings through contrast-enhanced computed tomography (CECT), magnetic resonance imaging (MRI), or transabdominal ultrasonography (137).

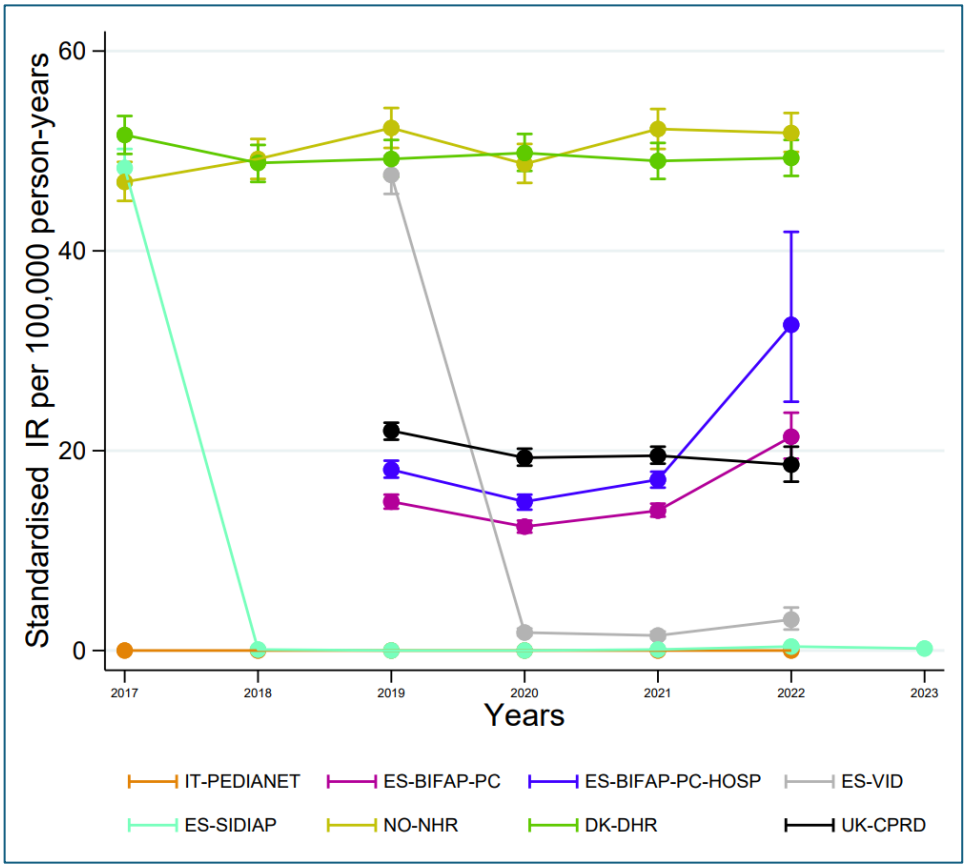
From a published extensive review of acute pancreatitis that reviewed 24 European studies, three Danish studies from 1980 to 2012 reported mean annual IRs ranging from 22.8 to 32.0/100000 PY, whereas a single study in Norway reported mean annual IRs of 14.7/100000 PY. UK IRs ranged from 17.6 to 93.9/100000 PY from 6 different studies covering different study periods across 1961 to 2010. Western Europe IRs ranged from 31.3/100000 PY in Italy (2000-2007) to 59.4/100000 PY in Spain (2001-2011) in two different studies (138). Overall, the global age-standardized incidence rate is reported to be 34.8/100,000 in 2019 (139).

In our study, most of the standardized annual IRs are in line with the overall fluctuation of incidences reported in previous studies and the ones reported on the CVM study (22): rates (100,000 PY) of pancreatitis are 12.4 to 21.4 in BIFAP_PC between 2019-2022, 14.9 to 18.1 in BIFAP_PC_HOSP between 2019-2021 with a peak of 32.6 in 2022 that would need further investigation (it may be due to the presence of data from only 1 Spanish region out of 5 compared to the other years), and 18.6 to 23.9 in CPRD between 2018-2022 which are also comparable to what we know from incidences reported in UK (137). NHR shows stable rates across the study period from 46.9 to 52.2, these were not previously detected in the CVM study. Similar stable rates are observed for DHR during the study period, ranging from 48.8 to 51.6. Rates are different from CVM values for SIDIAP, which report high IRs (100,000 PY) in 2017 of 48.34 and very low rates of <0.5 between 2018-2023, as well as VID that shows high incidence in 2019 (47.6), as in CVM in 2019 (47.6), and very low values from 2020 to 2022 (1.5 to 3.1). This pattern in these two data sources needs to be clarified.

Table 38 Standardized incidence rates of acute pancreatitis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				48.3	46.9	51.6	0.0
2018	0.0				0.1	49.2	48.8	23.9
2019	0.0	14.9	18.1	47.6	0.0	52.3	49.2	22.0
2020	0.0	12.4	14.9	1.8	0.0	48.7	49.8	19.3
2021	0.0	14.0	17.1	1.5	0.1	52.2	49.0	19.5
2022	0.0	21.4	32.6	3.1	0.4	51.8	49.3	18.6
2023					0.2			

Figure 43. Standardized incidence rates of acute pancreatitis per data source.



10.1.4.25 Pericarditis

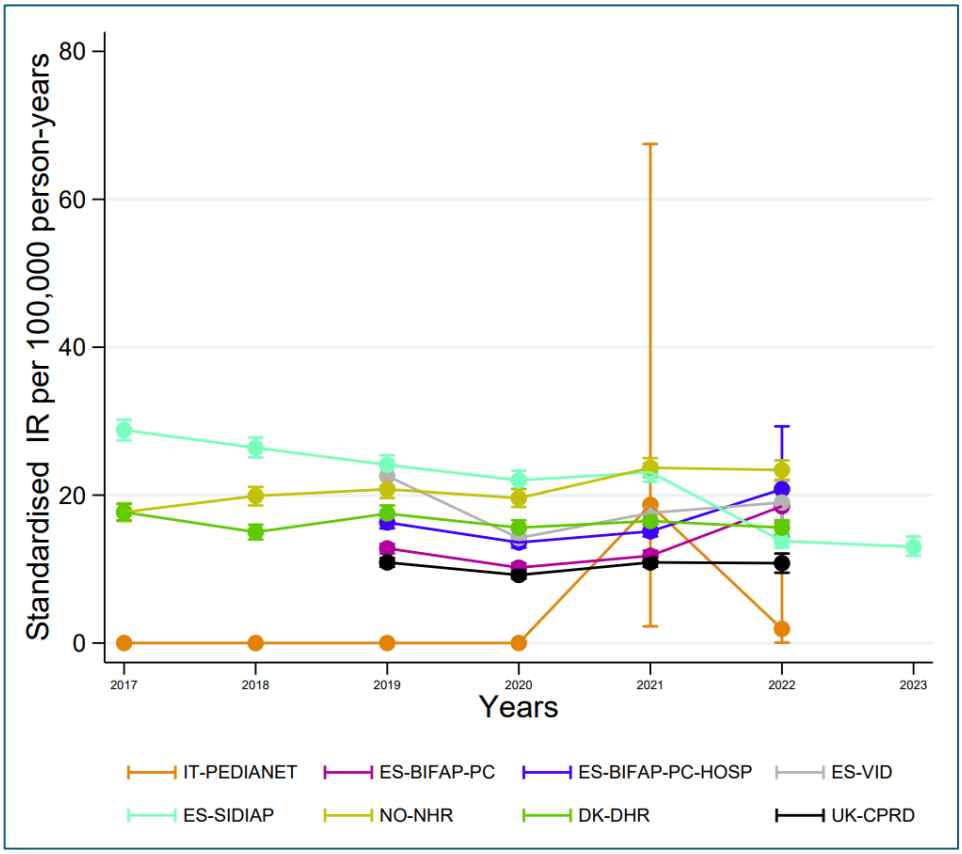
Pericarditis is an inflammation of the pericardium syndrome. It results in an increase in the normal volume of fluid surrounding the heart, usually leading to pericardial effusion or constrictive pericarditis. The aetiology and pathophysiology of pericarditis can be infectious (most commonly), or non-infectious, such as neoplasm, autoimmune process, injuries, or drug induced. Pericardium inflammation disease can be acute or chronic. Clinically, it is suggested by a characteristic chest pain description and the presence of a pericardial friction rub on auscultation. Electrocardiogram (ECG) and echocardiography are needed to confirm the diagnosis (140). The recurrence of the disease is estimated to occur in 15-30% of cases (141).

The reported incidence of acute pericarditis is approximately 27.7/100000 PY (142). In this study, standardized annual IRs are comparable to this rate as well as to the ones obtained in the CVM study (22). BIFAP_PC has IRs ranging from 10.2 to 18.5/100,000 PY between 2019 and 2022. BIFAP_PC_HOSP IRs range from 13.6 to 21.0/100,000 PY between 2019 and 2022. VID shows rates from 14.2 to 22.6/100,000 PY between 2019 and 2022. SIDIAP has higher rates from 2017 to 2021, ranging from 22.0 to 28.8/100,000 PY, and 13.8 and 13.0/100,000 PY in 2022 and 2023, respectively. NHR shows IRs from 23.7 to 17.7/100,000 PY between 2017 and 2022. CPRD IRs are 9.2-11.2/100,000 PY between 2018-2022. IRs in DHR are from 15 to 17.7/1000,000 PY between 2017 and 2022. PEDIANET shows no incidence from 2017 to 2020, with an increase to 18.7/100,000 PY in 2021 that would need to be further investigated. On average, databases containing hospital and emergency room data show higher IRs of pericarditis, in line with the reported IRs of 27.7/100000 PY.

Table 39 Standardized incidence rates of pericarditis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				28.8	17.7	17.7	0.0
2018	0.0				26.4	19.9	15.0	11.2
2019	0.0	12.8	16.3	22.6	24.1	20.8	17.5	10.9
2020	0.0	10.2	13.6	14.2	22.0	19.6	15.6	9.2
2021	18.7	11.8	15.1	17.6	23.1	23.7	16.5	10.9
2022	1.9	18.5	20.8	19.0	13.8	23.4	15.6	10.8
2023					13.0			

Figure 44 Standardized incidence rates of pericarditis per data source.



10.1.4.26 Polyarteritis nodosa (PAN)

Polyarteritis nodosa (PAN) is a rare necrotizing vasculitis of medium-sized arteries without glomerulonephritis or vasculitis in arterioles, capillaries, or venules, and not associated with ANCA. Vascular inflammation leads to microaneurysm formation, aneurysmal rupture with haemorrhage, thrombosis, and, consequently, organ ischaemia or infarction (143).

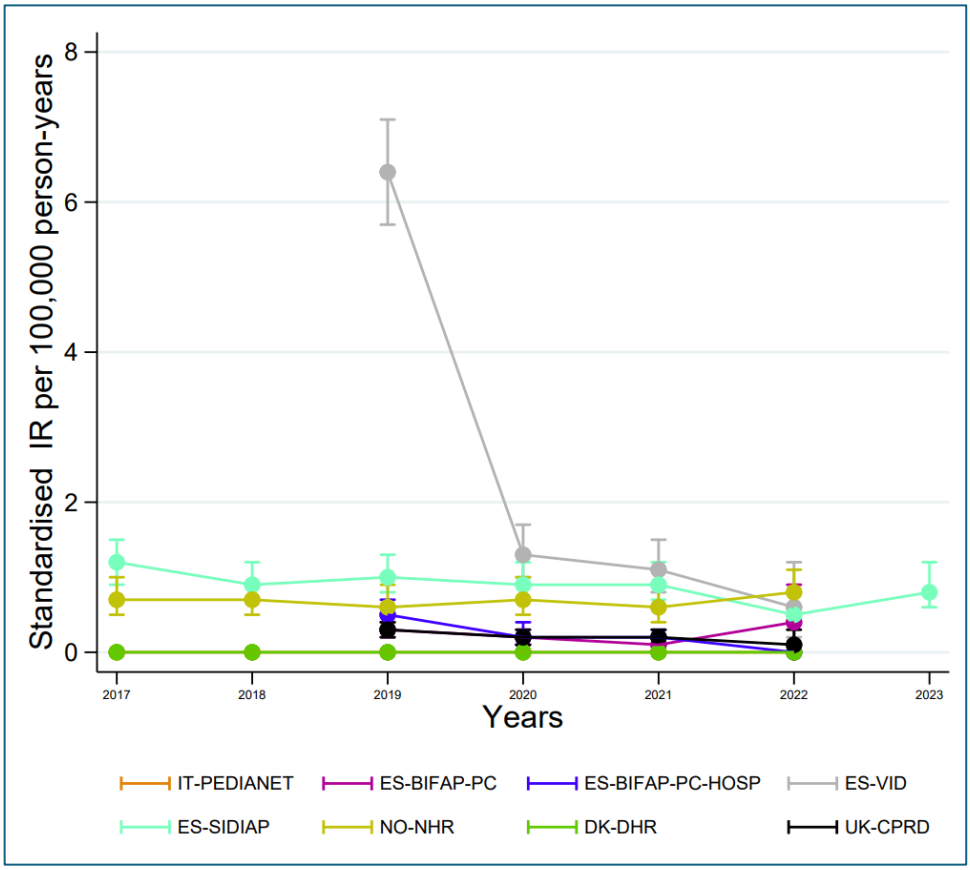
The annual European incidence of PAN is reported to be from 0 to 1/100,000 PY (144,145). Another French study reported a slightly higher incidence of around 3/100,000 PY (146).

In line with already published PAN rates, all data sources in this report show annual IRs lower or around 1/100,000 PY for PAN during the study period, except for VID in 2019 (6.4/100,000 PY). No data are available from DHR probably due to the presence of mainly hospital and emergency room data.

Table 40 Standardized incidence rates of polyarteritis nodosa per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				1.2	0.7	0.0	0.0
2018	0.0				0.9	0.7	0.0	0.6
2019	0.0	0.3	0.5	6.4	1.0	0.6	0.0	0.3
2020	0.0	0.2	0.2	1.3	0.9	0.7	0.0	0.2
2021	0.0	0.1	0.2	1.1	0.9	0.6	0.0	0.2
2022	0.0	0.4	0.0	0.6	0.5	0.8	0.0	0.1
2023					0.8			

Figure 45 Standardized incidence rates of polyarteritis nodosa per data source.



10.1.4.27 Psoriatic arthropathies (PsA)

Psoriatic arthropathies, or psoriatic arthritis (PsA), is a type of inflammatory arthritis often associated with psoriasis, whose manifestations are peripheral arthritis, dactylitis, enthesitis and axial involvement. It is characterized by the presence of HLA-B27-associated spondyloarthropathy, and the absence of rheumatoid factor. It is estimated to affect 30% of patients with psoriasis (147). In childhood, juvenile idiopathic arthritis (JIA) is the most common rheumatic disease (148) and skin psoriasis can be lacking.

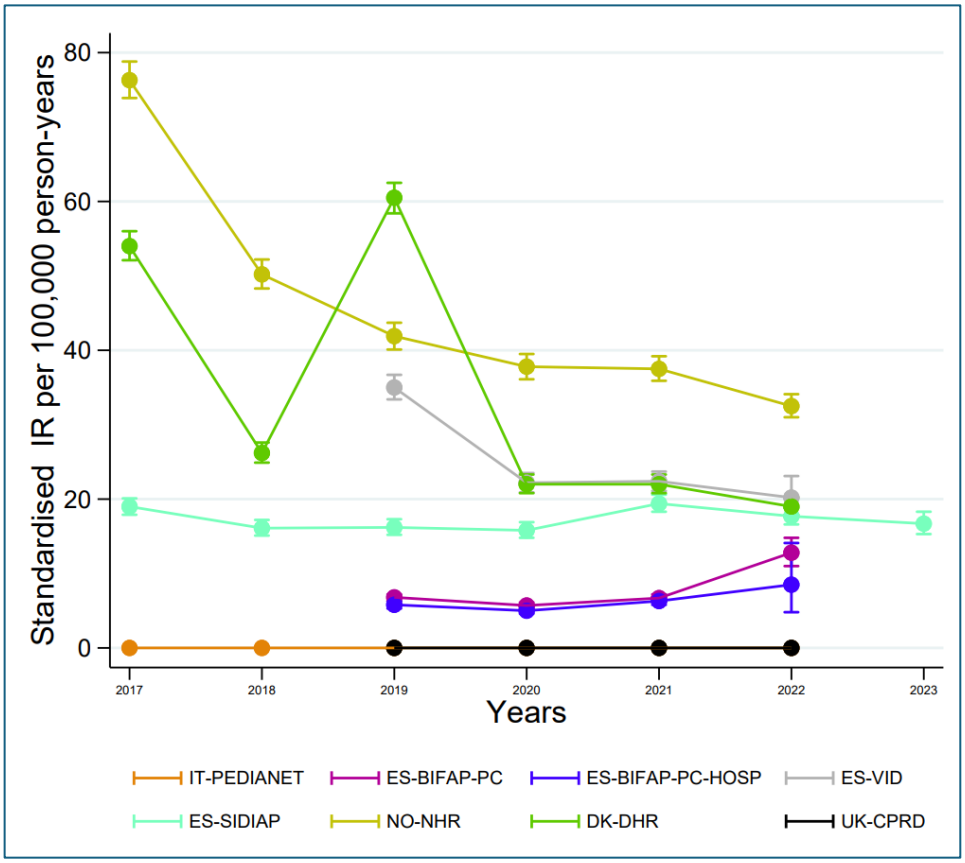
PsA incidence estimates of 2.0 to 8.0/100,000 PY are reported from most European countries (149,150). However, the incidence of PsA in the general population has been examined by relatively few studies. Recent publications range from 3.6–7.2/100,000 PY (151). Publications between 2001–2003 reported a much wider incidence range ranging from 0.1–23.1/100,000 PY (146).

In our study, BIFAP_PC_HOSP, SIDIAP, BIFAP_PC, and DHR shows IR values that are overall in line with reported European estimates and incidence ranges. SIDIAP shows IRs ranging from 16.2 to 19.4/100,000 PY during the study period (2017-2023), in line with 2001-2003 reported estimates. Slightly higher values are reported by VID (from 20.2 to 35.0/100000 PY), NHR (from 32.5 to 76.3/100000 PY), and DHR (from 19.0 to 60.5/100000 PY) during the study period. No values are reported in PEDIANET-IT and CPRD-UK.

Table 41 Standardized incidence rates of psoriatic arthropathies per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				19.0	76.3	54.0	0.0
2018	0.0				16.1	50.2	26.2	0.0
2019	0.0	6.8	5.8	35.0	16.2	41.9	60.5	0.0
2020	0.0	5.7	5.0	22.2	15.8	37.8	22.0	0.0
2021	0.0	6.7	6.3	22.4	19.4	37.5	22.0	0.0
2022	0.0	12.8	8.5	20.2	17.7	32.5	19.0	0.0
2023					16.7			

Figure 46 Standardized incidence rates of psoriatic arthropathies per data source.



10.1.4.28 Pulmonary embolism (PE)

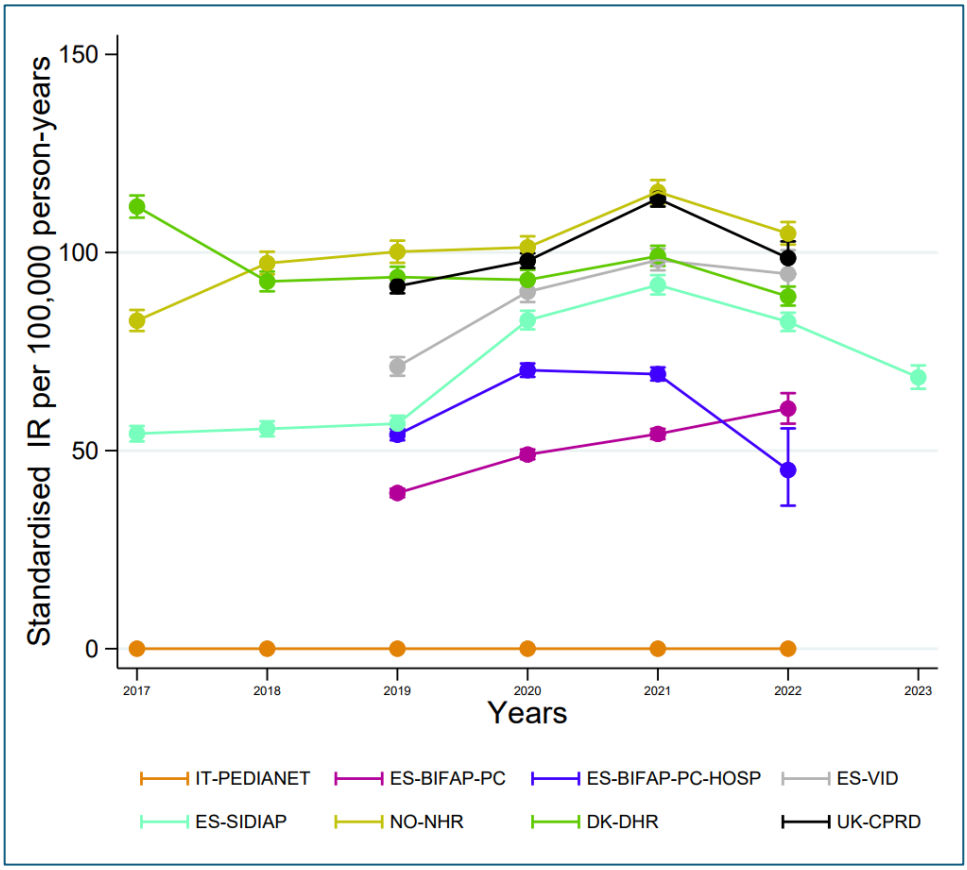
A PE usually occurs when a blood clot in a deep vein in the leg or pelvis breaks loose and travels through the blood to the lungs. It may not cause any signs or symptoms, or it may cause shortness of breath, chest pain, a bloody cough, or a fast or abnormal heartbeat. A pulmonary embolism can be life-threatening, especially if the blood clot is large or there are many clots. It may be caused by surgery and by certain medical conditions, such as cancer, heart disease, or lung disease. It can also be caused by taking medicines that contain hormones, pregnancy and childbirth, not moving for long periods of time, older age, smoking, obesity, and having a family history of blood clots or pulmonary embolism (152). While no exact epidemiological data are available, it is reported that PE has an annual incidence of 60 to 269 cases per 100,000 individuals in Europe and North America, with no evidence of variation between these two regions (153,154). The European guidelines for the diagnosis and management of PE and venous thrombosis report annual incidence rates of around 50-100/100000 PY (155).

In our study, all participating data sources reported data slightly below (BIFAP-ES data sources) or within this already reported window of incidence across the whole study period, except for PEDIANET that has no cases identified due to its paediatric sourcepopulation. In BIFAP_PC IRs were between 39.3 to 60.6/100,000 PY between 2019-2022 study period. BIFAP_PC_HOSP IRs were between 45.1 to 70.3/100,000 PY between 2019-2022 study period. VID shows IRs from 71.2 to 98.1/100,000 PY between 2019-2022 years. SIDIAP has IRs ranging from 54.3 to 91.8/100,000 PY between 2017-2023. NHR shows IRs from 82.8 to 115.3/100,000 PY between 2018-2022 years, in line with DHR observed estimates (88.9 to 111.6/100000 PY). CPRD has no cases in 2017, then incidences from 91.5 to 113.6/100,000 PY between 2018-2022. Again, all the observed rates are aligned with previously published estimations.

Table 42 Standardized incidence rates of pulmonary embolism per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				54.3	82.8	111.6	0.0
2018	0.0				55.5	97.3	92.7	97.4
2019	0.0	39.3	54.0	71.2	56.8	100.2	93.8	91.5
2020	0.0	49.0	70.3	90.1	82.9	101.3	93.1	97.9
2021	0.0	54.2	69.3	98.1	91.8	115.3	99.1	113.6
2022	0.0	60.6	45.1	94.6	82.5	104.8	88.9	98.6
2023					68.5			

Figure 47. Standardized incidence rates of pulmonary embolism per data source.



10.1.4.29 Rhabdomyolysis (RML)

Rhabdomyolysis (RML) is a condition generated by muscle cell injury and followed by the release of cell components into circulation, mainly proteins and electrolytes. RML can occur because of physical (i.e. trauma, excessive physical activity, vascular occlusion, sepsis, etc) or nonphysical causes (i.e. metabolic myopathies, medications, illicit drugs, endocrine disorders, etc) (156). It encompasses several symptoms and signs that include acute muscle weakness, myalgia, and muscle swelling combined with a creatine kinase cut-off value of > 1000 IU/L or > 5 × upper normal limit. Additionally, the substances released may cause acute kidney injury or heart damage, indicating a severe type of rhabdomyolysis (157).

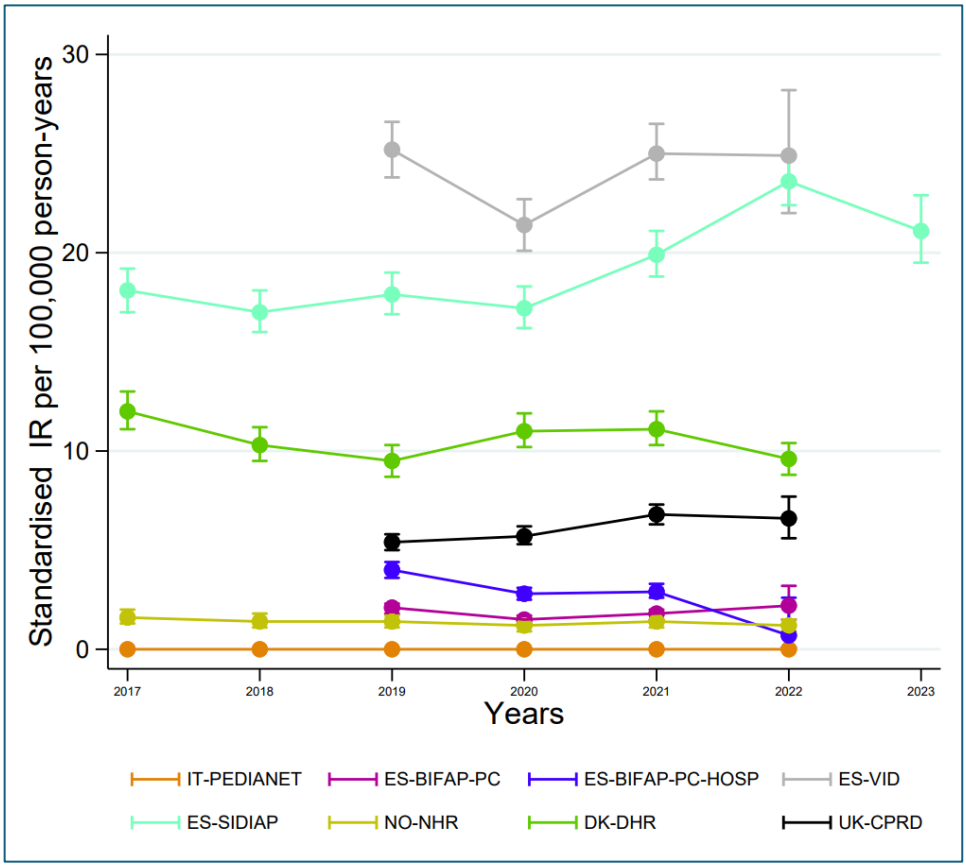
Approximately 26,000 cases of RML are reported annually in the United States (158). African Americans, males, obese patients, patients younger than ten years of age, and patients older than 60 years old all have a higher incidence of RML(159). However, the true incidence of RML is difficult to determine because of the lack of prospective studies and its wide etiology (160). Most of the studies are cause-specific (i.e. statins-induced RML (161), exertional RML, (162), etc.) which make it difficult to determine generalized annual IRs.

In our study, the annual IRs during the whole study period reported from BIFAP_PC_HOSP (2.8 to 4.2/100,000 PY), VID (21.4 to 25.2/100,000 PY), and SIDIAP (17.0 to 23.6/100,000 PY) are perfectly in line with the background rates of the CVM study (22) for 2019. BIFAP_PC and CPRD reported primary care-only data and the IRs are very low during the study period: 1.5 to 2.2 and 0 to 6.8/100,000 PY, respectively. This was also observed for the CVM rates. NHR has reported low IRs during the whole study period, ranging from 1.2 to 1.6/100,000 PY (163). This incidence observed in NHR is in line with another study in Norway that reported similar rates of 1.0 to 4.5/100000 PY within 2008-2014 (163). DHR rates, resulting from hospitalization and emergency rooms settings only, are slightly higher than primary care-only observed incidences (9.5 to 12.0/100000 PY). No cases are identified in PEDIANET.

Table 43 Standardized incidence rates of rhabdomyolysis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				18.1	1.6	12.0	0.0
2018	0.0				17.0	1.4	10.3	4.0
2019	0.0	2.1	4.0	25.2	17.9	1.4	9.5	5.4
2020	0.0	1.5	2.8	21.4	17.2	1.2	11.0	5.7
2021	0.0	1.8	2.9	25.0	19.9	1.4	11.1	6.8
2022	0.0	2.2	0.7	24.9	23.6	1.2	9.6	6.6
2023					21.1			

Figure 48. Standardized incidence rates of rhabdomyolysis per data source.



10.1.4.30 Rheumatoid arthritis (RA)

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterized by joint swelling, joint tenderness, and destruction of synovial joints, leading to severe disability and premature mortality. Autoimmunity and the overall systemic and articular inflammatory load drive the destructive progression of the disease (164–166). RA may start causing joint damage during the first year or two that a person has the disease. Once joint damage occurs, it generally cannot be reversed. RA can affect other organs, leading to monitoring for cardiovascular or respiratory health. The annual incidence of RA is reported to be about 20–50/100.000 PY in European countries (167–169). RA is reported to be more prevalent in women than men (lifetime risk of 3.6% and 1.7% respectively) (170). RA risk increases with age, with a peak incidence between 65 to 80 years old (171,172). In an Italian study, the yearly incidence of active RA per 100,000 PY is reported as 48 and 20 for women and men, respectively (173).

In this study, BIFAP_PC-HOSP, BIFAP_PC, and SIDIAP reported rates in line with the ones reported for European countries in other studies, with values ranging from 16.2 to 60.3/100.000PY across 2018/2019 to 2023, as well as NHR from 2018 to 2022 and CPRD and DHR in 2020, 2021, and 2022.

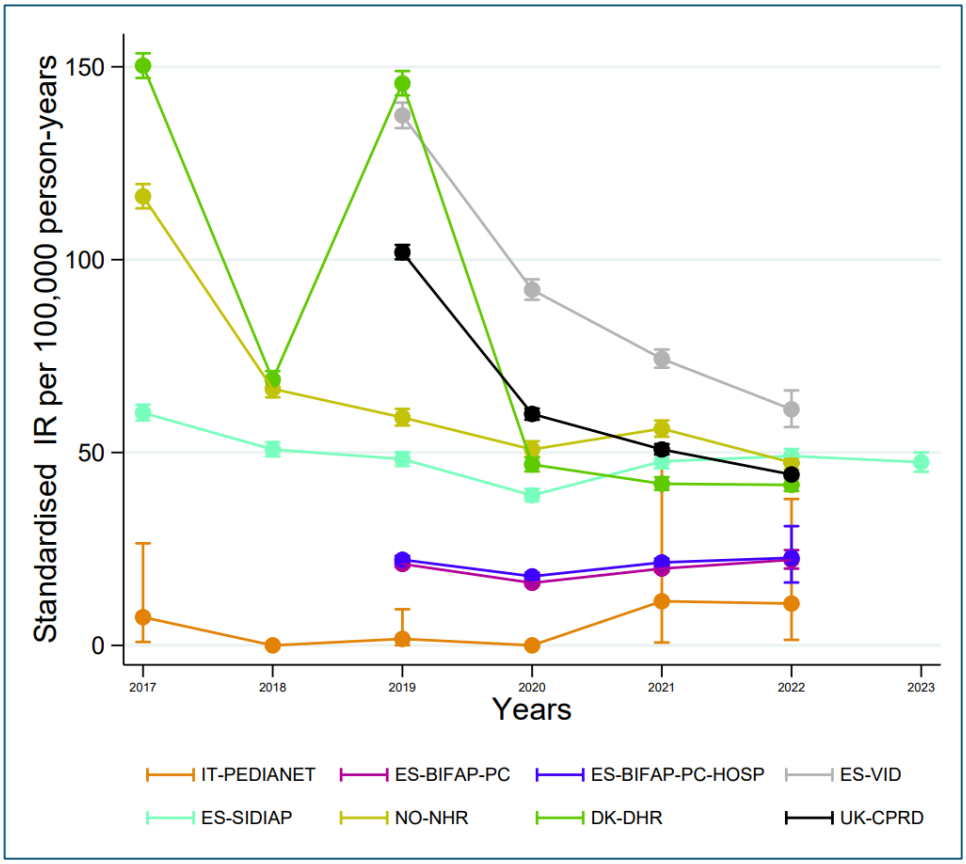
Higher IR values (100.000 PY) are reported from VID and would benefit from further investigation: 137.4 in 2019, 92.2 in 2020, 74.3 in 2021, and 61.2 in 2022. A similar situation is shown in CPRD for the years 2018 and 2019 with IRs of 243.3 and 101.9/100,000 PY, respectively.

SIDIAP showed a slightly higher incidence in 2017 of 60.3/100.000 PY (58.33; 62.42) as well as NHR in both 2017 and 2018 (116.4 and 66.5/100.000 PY, respectively) and DHR in 2017 (150.3) and 2019 (145.7). These increased rates may require further investigations to understand the reasoning behind. PEDIANET shows a low incidence of RA from 2017 to 2020, ranging from 0 to 7.32/100.000 PY. However, an increase to around 11/100.000 PY is found in 2021 and 2022 and would require a more in-depth assessment to be understood.

Table 44 Standardized incidence rates of rheumatoid arthritis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	7.3				60.3	116.4	150.3	0.0
2018	0.0				50.8	66.5	68.9	243.3
2019	1.7	21.1	22.2	137.4	48.3	59.1	145.7	101.9
2020	0.0	16.2	17.9	92.2	38.9	50.8	46.9	60.0
2021	11.4	19.9	21.5	74.3	47.7	56.2	41.9	50.8
2022	10.8	22.2	22.7	61.2	49.1	47.4	41.6	44.3
2023					47.5			

Figure 49 Standardized incidence rates of rheumatoid arthritis per data source.



10.1.4.31 Severe cutaneous adverse reactions to drug (SCARs)

SCARs include a broad spectrum of entities, mainly consisting of (174):

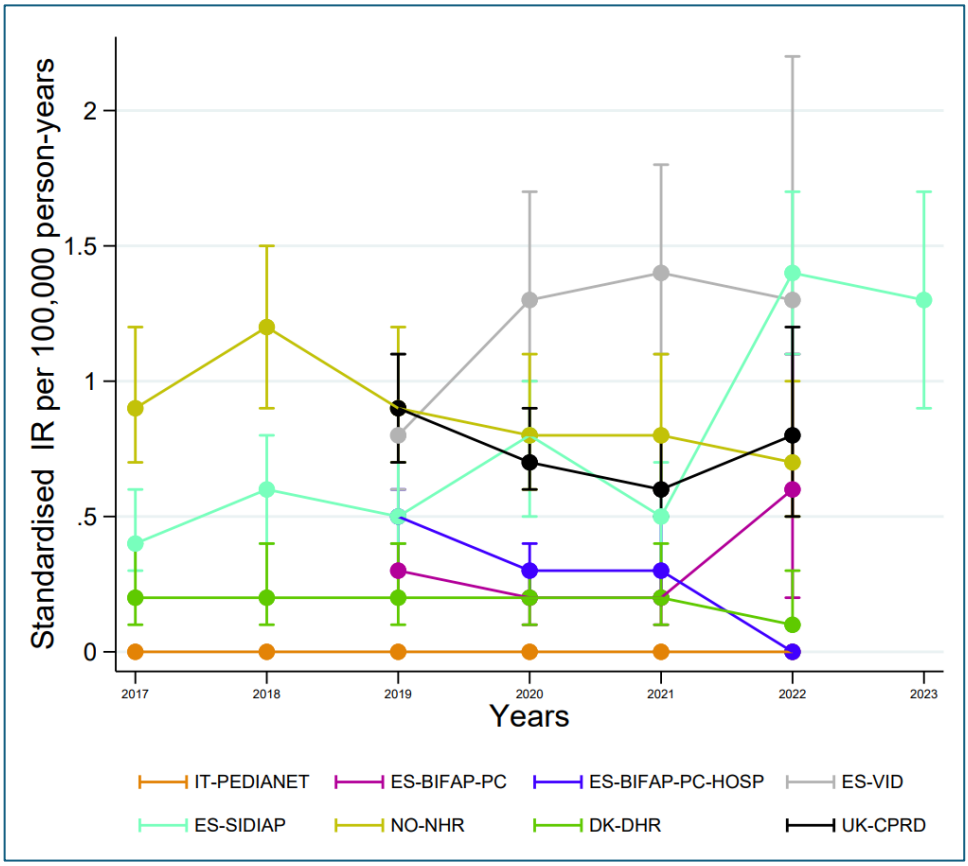
1. Stevens-Johnson syndrome (SJS) and Toxic epidermal necrolysis (TEN): both are variants of epidermal necrolysis. They occur 4–28 days after drug exposure. Disease is characterized by general physical deterioration, fever, and skin pain. SJS and TEN might be accompanied by lympho- and neutropenia, and renal impairment.
2. Drug reaction with eosinophilia and systemic symptoms (DRESS) syndrome: it usually begins 2–6 weeks after drug exposure. Clinical dermatological symptoms consist of facial oedema, erythroderma, distal oedema, purpura, pustules, and sometimes mucosal involvement. DRESS is accompanied by significant eosinophilia.
3. Acute generalised exanthematous pustulosis (AGEP): its onset is 2–11 days after drug exposure. Cutaneous symptoms develop simultaneously with high fever and numerous small, primarily non-follicular sterile pustules, arising on large areas of oedematous erythema in the major intertriginous zones.

In this study, we observed that IRs are similar to the ones reported in the CVM study (22) and across countries during the whole study period. Overall, incidences ranges from 0 to 1.44/100,000 PY. VID and SIDIAP reported slightly higher incidences in the year post pandemic (2020 to 2023). PEDIANET did not detected any cases. A recent review of background rates of SCARs for the safety assessment of COVID-19 in the US population reported incidences from 0.53 to 6.3/100,000 persons for SJS, from 0.04 to 0.5/100,000 persons for TEN, and from 0.08 to 0.16/100,000 persons for SJS/TEN (116). Although comparable, these US IRs are a slightly lower than the rates in this report, it might be due to the broader event definition used in this study (also including DRESS and AGEP).

Table 45 Standardized incidence rates of SCARS per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				0.4	0.9	0.2	0.0
2018	0.0				0.6	1.2	0.2	1.0
2019	0.0	0.3	0.5	0.8	0.5	0.9	0.2	0.9
2020	0.0	0.2	0.3	1.3	0.8	0.8	0.2	0.7
2021	0.0	0.2	0.3	1.4	0.5	0.8	0.2	0.6
2022	0.0	0.6	0.0	1.3	1.4	0.7	0.1	0.8
2023					1.3			

Figure 50 Standardized incidence rates of SCARS per data source.



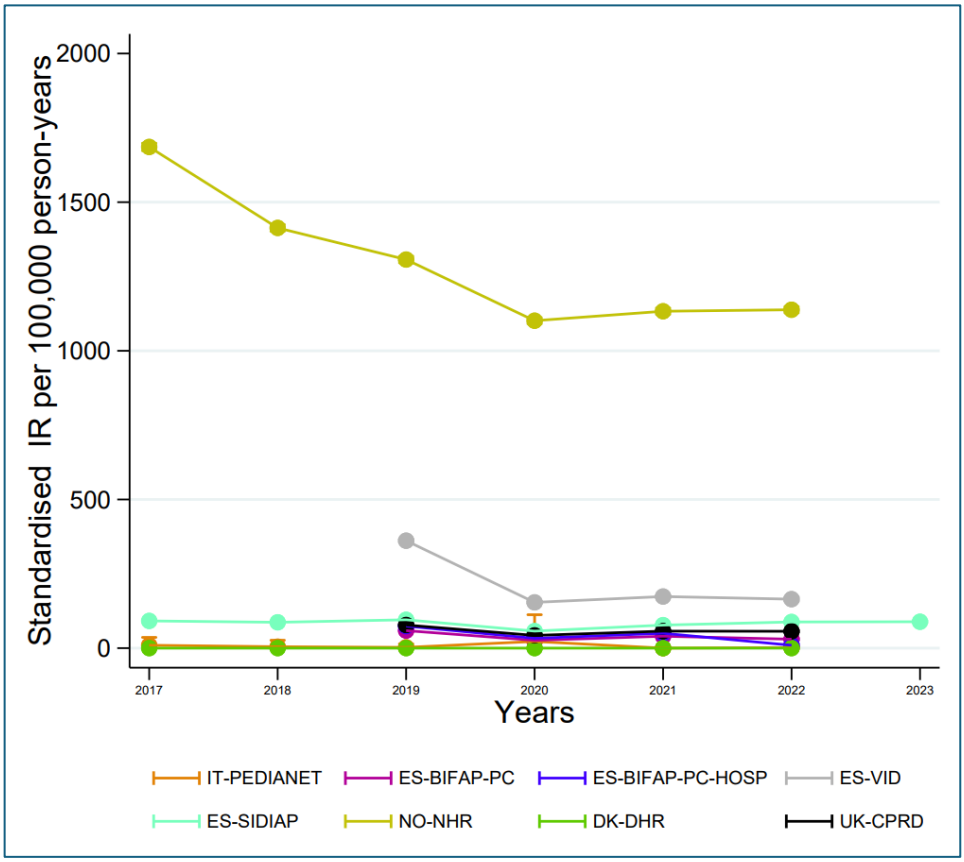
10.1.4.32 Sensorineural hearing loss (SNHL)

Hearing loss is the most prevalent sensory deficit and by the 5th leading cause of disability in adulthood, estimated to affect 6.1% of the world population by WHO. There are a few types of hearing loss, including conductive hearing loss (due to middle ear disease) and sensorineural hearing loss (SNHL). To properly identify the type of hearing loss it is essential to assess hearing with audiometry in an objectively way. To diagnose SNHL an audiometry must show hearing loss of ≥30 dB in three consecutive frequencies. Some risk factors are congenital or acquired infections, or ototoxic drugs, among others (175). In this study, SHNL IRs are well comparable for most of the data sources with the background rates reported in the CVM study prior-COVID-19 period (22). As shown in the CVM results, background annual rates in 2020 decreased probably due to the pandemic effects. After COVID-19 (2021 and 2022 mainly), BIFAP_PC, BIFAP_PC_HOSP, VID, SIDIAP, and CPRD maintained lower IRs than prior COVID-19 and this is different from what was observed in the CVM study, probably due to a longer and more complete follow-up period. It is important to underline that all the IRs reported by NHR are very high (>1000/100,000 PY) compared to the other participating data sources, CVM values, and other epidemiologic studies (176). By reviewing the code counts in **Annex 7**, we have found that in NHR database an important number of cases are identified through the ICD10 code H91.1, referring to presbycusis, a bilateral age-related hearing loss do not recommend for the identification of SHNL by the SPEAC AESI case definition companion guide (176). Moreover, the use of this diagnosis code could negatively impact the incidence rates in all data sources. PEDIANET has lower rates (from 1.7 to 22.4/100,000 PY) of SHNL during the study period compared to other countries, and this is in line with the age-dependency of the associated risk, with a peak of 22.4/100,000 PY in 2020 during the pandemic that may benefit from further inspection. Finally, the SPEAC companion guide on sensorineural hearing loss reports incidence rates from 7 epidemiologic studies, none of them in European population (176). In the US, the incidence rates range from 10.7 to 27/100,000 PY. In Asian population, it goes from 14 to 65/100,000 PY. The last are the most comparable to our results.

Table 46 Standardized incidence rates of sensorineural hearing loss per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	9.7				91.5	1685.7	0.0	0.0
2018	4.8				86.9	1413.4	0.0	76.3
2019	2.6	58.9	74.2	361.4	95.3	1307.0	0.0	77.7
2020	22.4	26.2	33.0	153.9	57.3	1101.3	0.0	42.8
2021	0.0	40.0	50.0	173.6	77.4	1133.1	0.0	57.2
2022	1.7	30.0	9.6	164.7	88.0	1138.2	0.0	56.7
2023					88.7			

Figure 51. Standardized incidence rates of sensorineural hearing loss per data source.



10.1.4.33 Single organ cutaneous vasculitis (SOCV)

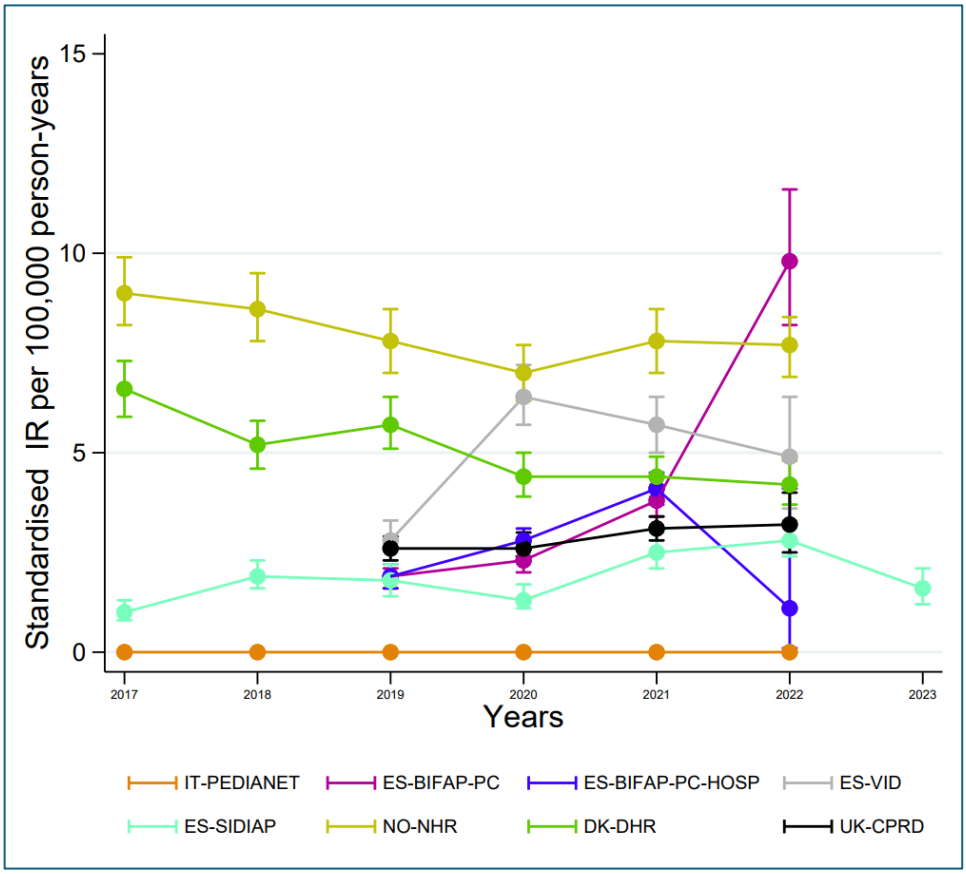
Single Organ Cutaneous Vasculitis (SOCV) is a syndrome with clinical and histological features of small vessel vasculitis in the skin. It can be the first clinical sign of systemic vasculitis. The main diagnostic procedure is a skin biopsy. Histology typically shows perivascular inflammatory cells infiltrate with leukocytoclasia, erythrocyte extravasation, or haemorrhage into the dermis and fibrinoid necrosis or degeneration of the dermal postcapillary venules. In 90% of patients SOCV will be resolved in weeks to months of onset. Only simple measures are recommended like bed rest with elevation of the lower limbs and treatment with nonsteroidal anti-inflammatory drugs or antihistamines (177). (136).

Cutaneous Vasculitis (CV) occurs in all age groups, has a slight female predominance, and is much more common in adults than in children (about 90% of cases in adults and 10% in children) (177). It is reported that the annual incidence of biopsy-proven CV of all types ranges from 1.5-6.0/100000 PY, however this incidence may be underestimated as patients with clinically obvious and/or mild disease may not have been biopsied, or their specimens were interpreted by another, private laboratory (178,179). Overall, all the SAFETY-VAC data sources reported SOCV incidences ranging from 1.6 to 8.6/100000 PY that are in line with general annual global estimations. Moreover, in this study, reported IRs are lower than the once reported in the ACCESS study (6), and this is in line with the trend observed in the CVM study (22). Compared with the CVM results, BIFAP_PC and BIFAP_PC_HOSP showed an annual IRs higher for the post-COVID-19 period (BIFAP_PC: 3.8 in 2021 and 9.8 in 2022; BIFAP_PC_HOSP: 4.1/100000 PY in 2021). VID and SIDIAP showed lower annual IRs than CVM background rates values during the whole study period. Only CPRD showed slightly higher values than CVM (but lower than ACCESS) across the study period. No events were found in PEDIANET, as expected with the increment of the incidence age-dependently, as shown in the CVM study. Rates reported by Pottegard et al. (73) were 20 and 14/100,000 PY in Denmark and Norway respectively and are slightly higher than the ones reported here for DHR (4.2-5.7/100000 PY) and NHR (4.2-6.6/100000 PY). However, DHR and NHR rates herein reported are in line with global estimations.

Table 47 Standardized incidence rates of SOCV per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				1.0	9.0	6.6	0.0
2018	0.0				1.9	8.6	5.2	2.8
2019	0.0	1.9	1.9	2.8	1.8	7.8	5.7	2.6
2020	0.0	2.3	2.8	6.4	1.3	7.0	4.4	2.6
2021	0.0	3.8	4.1	5.7	2.5	7.8	4.4	3.1
2022	0.0	9.8	1.1	4.9	2.8	7.7	4.2	3.2
2023					1.6			

Figure 52 Standardized incidence rates of SOCV per data source.



10.1.4.34 Systemic lupus erythematosus (SLE)

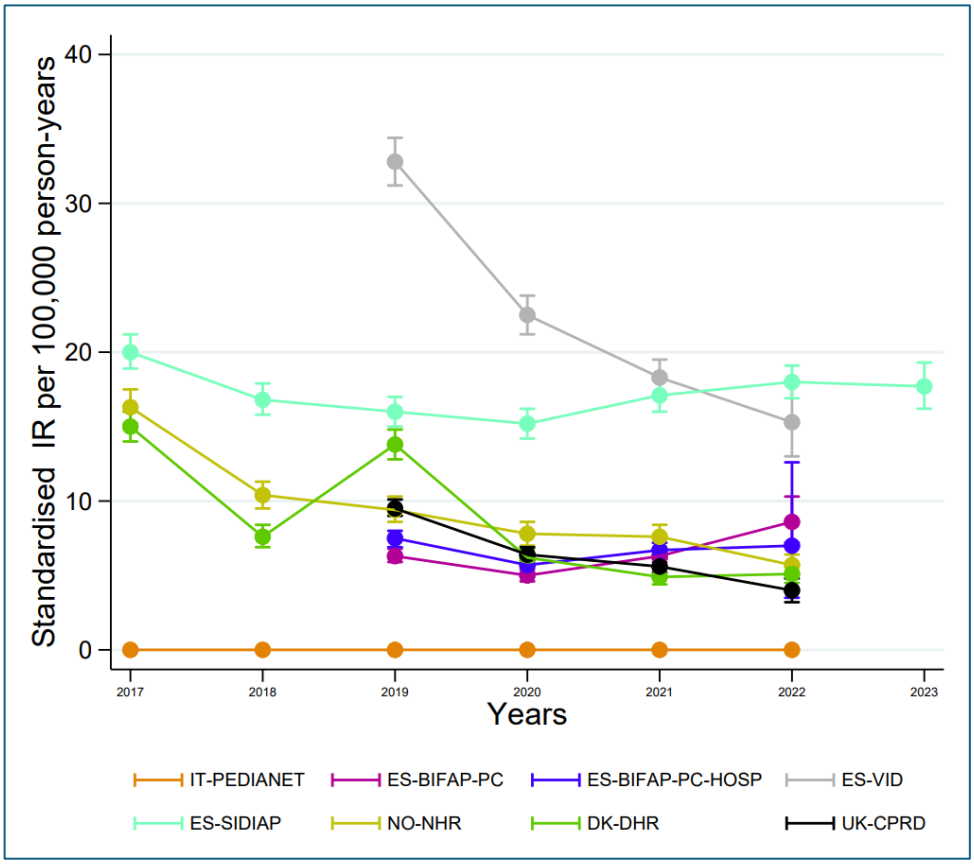
Systemic lupus erythematosus (SLE) is a chronic, relapsing, inflammatory, and often febrile multisystemic disorder of connective tissue, characterized principally by involvement of the skin, joints, kidneys, and serosal membranes. It is of unknown etiology but is thought to represent a failure of the regulatory mechanisms of the autoimmune system. The disease is marked by a wide range of system dysfunctions, an elevated erythrocyte sedimentation rate, and the formation of LE cells in the blood or bone marrow (180,181).

The global SLE incidence and newly diagnosed population are estimated to be 5.14 (1.4 to 15.13) per 100,000 PY, specifically, the incidence in Europe is reported to be 0.3 to 5.1/100,000 PY (182,183). In our study, BIFAP_PC-HOSP (primary diagnosis only), BIFAP_PC, and NHR reported rates in line with the global and European values, whereas SIDIAP has higher rates during the 2017-2023 period (from 15.9 to 20.0/100,000 PY) as well as VID from 2020-2023 (15.3 to 22.5/100,000 PY), both including primary, secondary and specialist diagnosis. VID also shows a strong incidence increment in 2019 (32.7/100,000 PY). No IRs are reported for PEDIANET. CPRD shows IRs in line with the reference values for 2020, 2021, and 2022, whereas slightly higher IRs are observed in 2018 (18.2/100,000 PY) and 2019 (9.5/100,000 PY). DHR incidence rates in this study are similar to the 6.9/100,000 PY reported as crude incidence rate by Willame C., et al (50). Moreover, the same author reports a pooled crude SLE rate of 5.3/100,000 PY using European electronic healthcare record databases.

Table 48 Standardized incidence rates of SLE per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				20.0	16.3	15.0	0.0
2018	0.0				16.8	10.4	7.6	18.3
2019	0.0	6.3	7.5	32.8	16.0	9.4	13.8	9.5
2020	0.0	5.0	5.7	22.5	15.2	7.8	6.2	6.4
2021	0.0	6.3	6.7	18.3	17.1	7.6	4.9	5.6
2022	0.0	8.6	7.0	15.3	18.0	5.7	5.1	4.0
2023					17.7			

Figure 53 Standardized incidence rates of SLE per data source.



10.1.4.35 Thrombocytopenia (TP)

Thrombocytopenia (TP) refers to an abnormally low platelet count (generally less than 150x10^9/L). Insufficient production, abnormal distribution, or excessive destruction of platelets define pathogenic mechanisms. Excessive destruction can be caused by microangiopathy, hereditary platelet abnormalities, or immunologic mechanisms. Immunologic TP can be caused by autoimmune mechanisms, neonatal isoimmunization, or a nonspecific immune response. Idiopathic TP (ITP) refers to TP without an identified aetiology, although an autoimmune aetiology is frequently suspected but not always verified through exhaustive exclusion of differential diagnoses. It is usually related to the presence of clinical signs and symptoms of spontaneous bleeding (184).

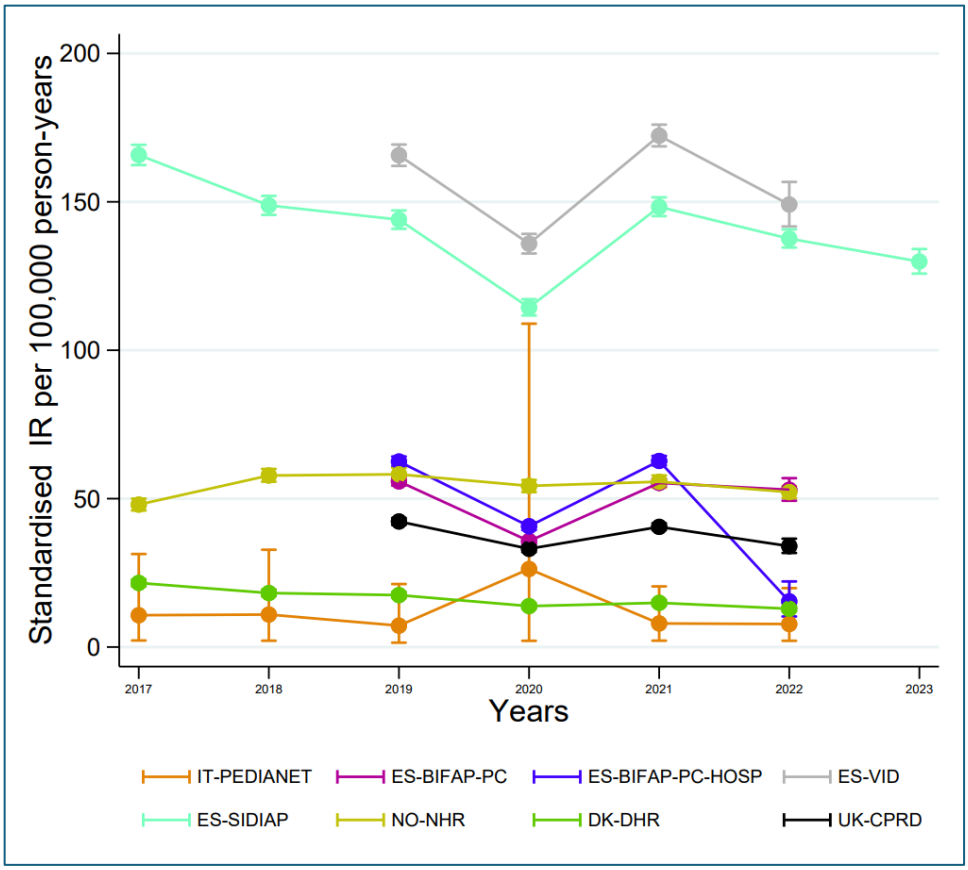
Incidence rates (IRs) reported in the literature for thrombocytopenia events vary according to the definitions used. Some authors include both immune and secondary thrombocytopenia in their definitions, while others include only one. When both immune and secondary thrombocytopenia are included (6), the IR is reported to be 92.4/100000 PYs in females and 147.2/100000 PY in males between 2017 and 2019. Additionally, the type of data source affects the reported IR: a higher IR is observed when combining data from general practitioners (GPs) with hospital data (both in- and out-patient settings), at 92.1/100,000 PY, compared to in-patient data only, which shows an IR of 18.0/100000 PYs (6). The reported IR of immune thrombocytopenia during the pre-pandemic period (2017-2019) ranges from 8 to 85/100000 PY (185). Notably, higher IRs are reported when data sources include both hospital and primary care (11-85/100000 PY (185), 8-56/100000 PYs (69), compared to in- and out-patient settings with emergency room data (18-26/100000 PYs (186). Similar to the first study, the IR is slightly higher in males compared to females, particularly in older age groups (69,185). Another study assessing IRs and 2014, reported an IR of 21.7/100000 PY using a broad definition of the event, as observed in GP and outpatient settings (50). Finally, Nasreen S. et al. reported an IR of idiopathic thrombocytopenia of 43.9/100,000 PY between 2015 and 2019 using administrative databases for hospitalizations and emergency department visits (115). Notably, the mean IR was higher in males than females over 40 years of age. Overall, the IRs observed from SAFETY-VAC data sources are mostly in line with already published evidence. In this study, we observed annual IRs for VID, which include both primary care and hospital information, that are almost stable across years and comparable with the reported CVM values (22) for 2019 (165.7/100000 PY). BIFAP-PC reported IRs stable across the whole study period (35.7 to 58.0/100000 PY) in line with ACESS primary care-only data but much lower than background rates reported in the CVM study. BIFAP_PC_HOSP (primary diagnosis only) has IRs ranging from 15.5 to 65.6/100,000 PY across the study period and overall lower than both CVM (22) and ACCESS values (6). However, those rates are in line with published literature. SIDIAP IRs ranges from 114.4 to 165.8/100000 PY and are in line with CVM background rates but lower than ACCESS IRs, as well as with studies including primary care and hospital data. PEDIANET-IT reported IRs lower than 11.0/100,000 PY, comparable with CVM values. NHR shows comparable IRs across years of the study (48.5 to 58.2/100,000 PY) with ACCESS reported IRs. DHR rates are stable during the study period and ranges from 12.9 to 21,6/100000 PY. Overall, our rates cannot be compared with the rate reported by Li et al, which vary across data source between 1-100/100,000 PY (68). Thrombocytopenia rates presented are comparable to the rates reported in previous

studies in European and North American population, as for the CVM results (22). As mentioned in the CVM analyses, care needs to be paid to inclusion of secondary thrombocytopenia, which increases a lot with age.

Table 49 Standardized incidence rates of thrombocytopenia per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	10.7				165.8	48.0	21.6	0.0
2018	10.9				148.8	57.8	18.2	49.2
2019	7.2	55.8	62.5	165.7	144.0	58.2	17.5	42.3
2020	26.3	35.7	40.7	135.9	114.4	54.3	13.8	33.1
2021	7.9	55.3	62.7	172.3	148.3	55.7	14.9	40.5
2022	7.8	53.0	15.4	149.1	137.6	52.2	12.9	34.0
2023					129.9			

Figure 54 Standardized incidence rates of thrombocytopenia per data source.



10.1.4.36 Transverse myelitis

Transverse myelitis is a disorder caused by an inflammation of the spinal cord. The term transverse refers to the pattern of sensitivity changes— there is often a band-like sensation across the trunk of the body, with sensory changes below. Although some people recover from transverse myelitis with minor or no residual problems, the healing process may take months to years. There is no cure for transverse myelitis, but there are treatments to prevent or minimize permanent neurological deficits (187).

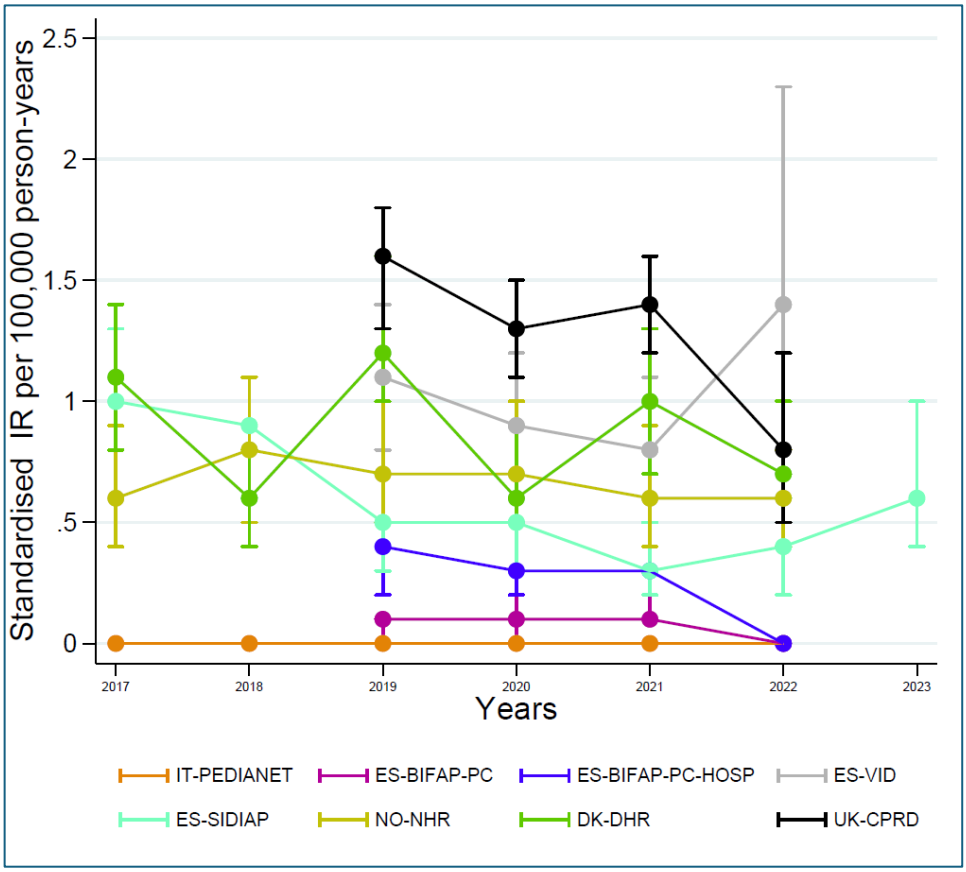
The background incidence rate of transverse myelitis reported by Li and colleagues (69) ranged from 1 to 4/100000 PY depending on age and sex strata, this estimate was obtained from subjects enrolled between 2017 and 2019 from 13 data sources, 8 collecting electronic health records and 5 comprising administrative claims, from eight different countries. Another multi-database study (185) involving 26 data sources (i.e., 8 administrative claims databases, 12 electronic health records, 1 electronic health records with a registry, and 5 general practitioner databases) from 11 countries reported a pooled incidence rate estimated in the pre-pandemic general population that ranged from 2 to 7/100000 PY, depending on age and sex strata. Another study (115) based on hospitalization and emergency room data from Ontario reported a mean annual rate observed during 2015–2019 of 0.8/100000 PY and 1.7/100000 PY using a narrow and a broad case definition, respectively. Results from the three above mentioned studies (69,115,185) suggest that the IR of transverse myelitis might slightly increase with age and female sex, however, no clear conclusion can be drawn also due to the very low frequency of the event.

Overall, the estimation of annual IRs of transverse myelitis in SAFETY-VAC data sources is aligned with already published evidence. The observed IRs in BIFAP-ES, VID-ES, and SIDIAP-ES are stable across the study period and in line with the CVM reported values (22). DHR rates are in line with literature reports and range from 0.6 to 1.1. NHR reported IRs slightly higher (0.6 to 0.8/100,000 PY across the study period), but overall comparable with the ACCESS project (6) and other publications (115). CPRD shows higher IRs (from 0.8 to 2.3/100,000 PY) than primary care-only results reported in ACCESS (0.4/100,000 PY) across the whole study period. No cases are identified in PEDIANET.

Table 50 Standardized incidence rates of transverse myelitis per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				1.0	0.6	1.1	0.0
2018	0.0				0.9	0.8	0.6	2.3
2019	0.0	0.1	0.4	1.1	0.5	0.7	1.2	1.6
2020	0.0	0.1	0.3	0.9	0.5	0.7	0.6	1.3
2021	0.0	0.1	0.3	0.8	0.3	0.6	1.0	1.4
2022	0.0	0.0	0.0	1.4	0.4	0.6	0.7	0.8
2023					0.6			

Figure 55 Standardized incidence rates of transverse myelitis per data source.



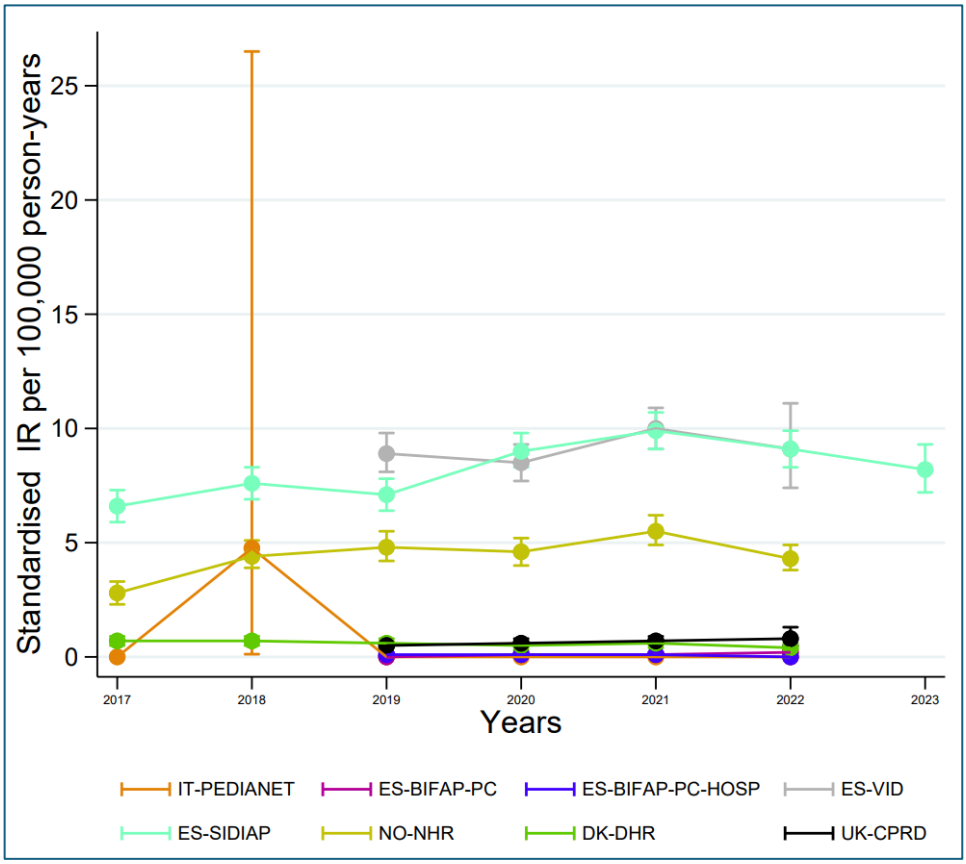
10.1.4.37 TTS

The algorithm utilized in this study to define TTS was the presence of thrombocytopenia plus any of the following events: deep venous thrombosis, pulmonary embolism, ischemic stroke, splanchnic venous thrombosis, acute myocardial infarction, central venous sinus thrombosis, and VTE in other locations, occurring within 10 days, before or after, from each other. TTS rates in the study by Burn et al. did not provide an overall rate estimation, but separated the components, showing high variability across different data sources (188).TTS is a rate event, in this study we report an annual IRs <0.2/100,000 PY across the study period in BIFAP_PC and BIFAP_PC_HOSP, which are in line with the CVM reported rates (22). CPRD has slightly higher IRs from 0.4 (2018) to 0.8 (2022)/100,000 PY, but comparable to the CVM rates. Moreover, an industry-sponsored post-authorization safety study aimed to estimate the incidence rate of TTS in the UK using integrated healthcare databases (189) produced a similar pre-COVID-19 crude incidence rate of 0.42 (IC95% 0.36-0.48)/100,000 PY. VID and SIDIAP, with IRs of 6.9-8.1 and 5.0-8.6/100,000 PY across the years, are also in similar to CVM values. NHR reported IRs varying from 2.0 to 4.6/100,000 PY across study years. Several data sources reported a slight increase of TTS rates in 2021. PEDIANET did not retrieved events except for a peak of 4.8/100,000 PY in 2018 that would benefit from further investigation.

Table 51 Standardized incidence rates of TTS per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				6.6	2.8	0.7	0.0
2018	4.8				7.6	4.4	0.7	0.4
2019	0.0	0.0	0.1	8.9	7.1	4.8	0.6	0.5
2020	0.0	0.1	0.1	8.5	9.0	4.6	0.5	0.6
2021	0.0	0.1	0.1	10.0	9.9	5.5	0.6	0.7
2022	0.0	0.2	0.0	9.1	9.1	4.3	0.4	0.8
2023					8.2			

Figure 56 Standardized incidence rates of TTS per data source.



10.1.4.38 Ulcerative colitis (UC)

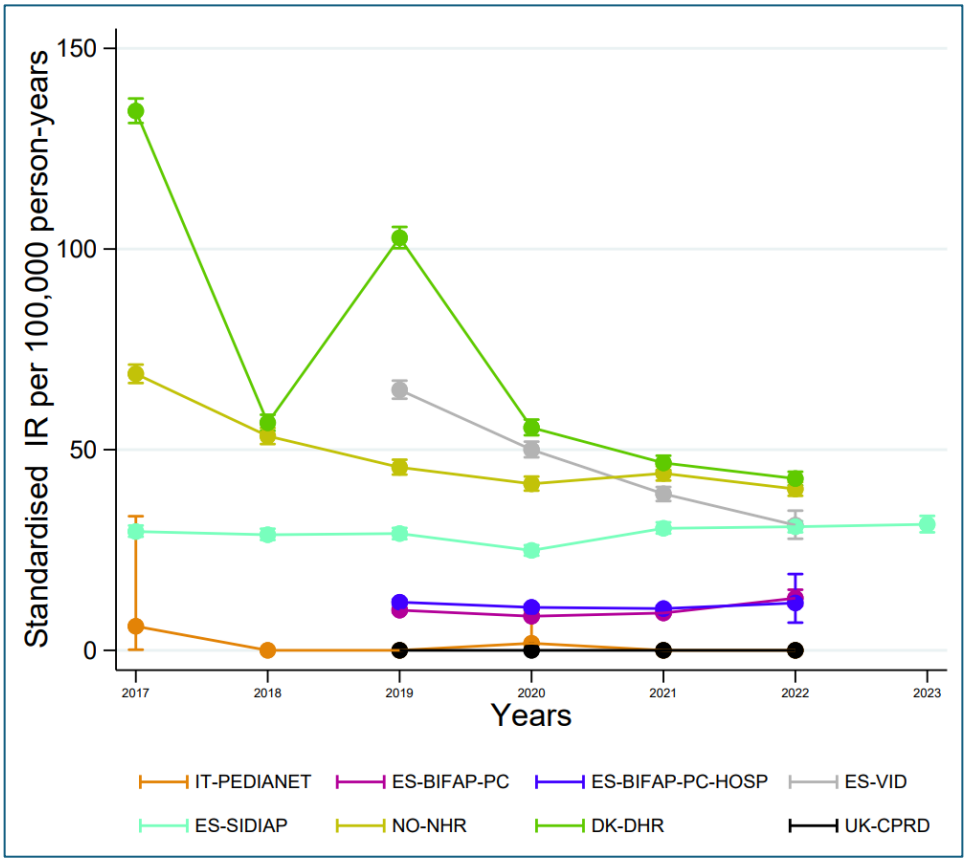
Ulcerative colitis (UC) is a type of inflammatory bowel disease (IBD) that is characterized by continuous and diffuse inflammation which is limited to the colonic mucosa and extends proximally from the rectum (190). It is considered a chronic immune-mediated inflammatory disorder of the colon that is hypothesized to be related to exposure to environmental risk factors leading to inappropriate immune responses to enteric commensal microbes in genetically susceptible individuals.

The highest incidences of UC have been reported in northern Europe (24.3/100000 PY) and annual IRs ranges from 1 to 17.2/100000 PY in the rest of Europe (191,192). In this study, we observed annual IRs in line with previously reported values in the European literature only for BIFAP_PC and BIFAP_PC_HOSP (8.5 to 13.0/100,000 PY). VID has higher values, with IRs ranging from 12.2 to 64.9 along the study period, as well as for SIDIAP (from 24.9 to 31.4/100,000 PY), NHR (from 40.1 to 68.9/100,000 PY) and DHR (from 42.8 to 134.4/100,000 PY). PEDIANET has no cases of UC except for 2020 (1.8/100,000 PY), possible related to SARS-CoV-2 infection (193). No data are available for CPRD in this data instance.

Table 52 Standardized incidence rates of ulcerative colitis per data source

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	6.0				29.6	68.9	134.4	
2018	0.0				28.8	53.4	56.7	
2019	0.0	10.0	12.0	64.9	29.1	45.6	102.8	
2020	1.8	8.5	10.7	50.0	24.9	41.5	55.5	
2021	0.0	9.3	10.4	39.0	30.4	44.1	46.7	
2022	0.0	13.0	11.8	31.2	30.8	40.2	42.8	
2023					31.4			

Figure 57 Standardized incidence rates of ulcerative colitis per data source.



10.1.4.39 Venous thromboembolism (VTE)

Venous thromboembolism is a condition that includes deep vein thrombosis (DVT) and pulmonary thromboembolism (PE). Deep vein thrombosis refers to the formation of a blood clot, called a thrombus, in one of the body's large veins. This formation often happens in the lower limbs, resulting in swelling and pain.

In Western countries, the annual incidence rate of VTE is reported to be approximately 100-200/100000 PY, translating to about 300,000-600,000 cases each year (194). In Eastern countries, the incidence tends to be lower, <100/100000PY (195). A study from Norway reported an incidence rate of VTE at 143/100000 PY, with deep vein thrombosis (DVT) at 93 and pulmonary embolism (PE) at 50/100000 PY (196). In a population-based cohort study of patients with COVID-19, the incidence rate of VTE for non-infected individuals was reported to be 237/100000 PY (197).

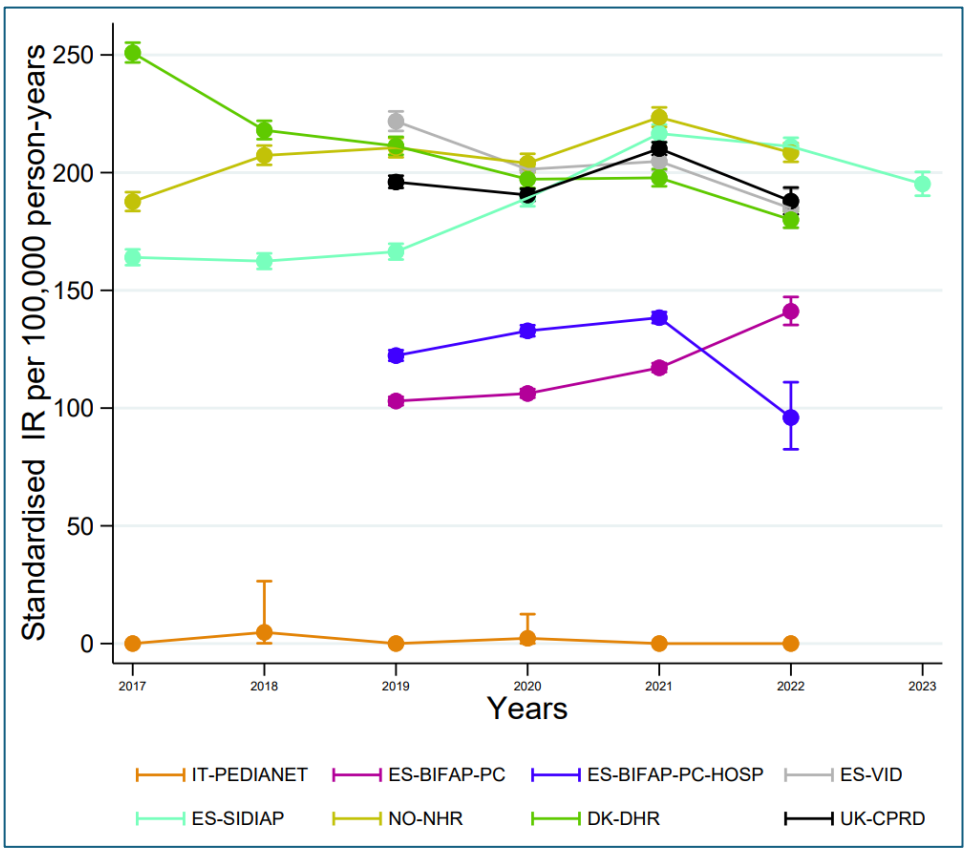
The risk of VTE according to sex varies by life stage. Women of child-bearing age are at increased risk of VTE, probably as a consequence of exposure to oral contraceptives and pregnancy. At older ages, men are at a modestly greater risk of VTE than women. The risk of VTE rises exponentially with age. For example, in the LITE study (198), the race-adjusted and sex-adjusted incidence of VTE per 100000 PY was 720 in those aged 40 to <75 years, 312 in those aged 75 to <85 years and 696 in those aged ≥85 years. In a more recent analysis, the hazard ratios of VTE per decade of age were 2.67 (95% CI 2.45–2.91) (in the Emerging Risk Factors Collaboration (n=731,728)) and 1.81 (95% CI 1.71–1.92) (in the UK Biobank (n = 421,537)) (199).

Overall, SAFETY-VAC incidences are in line with IRs estimation from Western countries. BIFAP_PC shows IRs a bit lower than the one reported by the ACCESS project (22) (141.8/100,000 PY), and this trend is in line with the CVM (6)for primary care-only data sources, with other studies by Gubernot et al. in the US (48), with general reported annual incidence rate of VTE mentioned above, and with results from Pottegård et al. in Norway and Denmark (72). Differently, CPRD has much higher rates during the study period (from 187.9 to 218.8/100000 PY) for being a GP-only data source. However, these values are in line with Western countries incidences. The IRs in BIFAP_PC_HOSP ranges from 96.0 to 138.4/100000 PY across the study period and are in line with the rates reported in the CVM study and reported global estimations. VID has rates in alignment with ACCESS and CVM. SIDIAP reported IRs comparable with both ACCESS and CVM values during the whole study period. NHR and DHR also show comparable and stable IRs that are also comparable with the ACCESS values for primary care and/or hospitalization-based data across the whole study period. As expected, almost no cases are detected in PEDIANET due to the presence of paediatric population only.

Table 53 Standardized incidence rates of venous thromboembolism per data source.

year	IT-PEDIANET	ES-BIFAP-PC	ES-BIFAP-PC-HOSP	ES-VID	ES-SIDIAP	NO-NHR	DK-DHR	UK-CPRD
2017	0.0				164.0	187.7	250.9	0.0
2018	4.8				162.4	207.3	218.0	218.8
2019	0.0	103.0	122.3	221.8	166.4	210.6	211.3	196.0
2020	2.2	106.2	132.8	201.4	189.2	204.0	197.2	190.5
2021	0.0	117.1	138.4	204.8	216.6	223.5	197.8	210.2
2022	0.0	141.1	96.0	184.8	211.1	208.5	180.0	187.9
2023					195.2			

Figure 58 Standardized incidence rates of venous thromboembolism per data source.



10.2 Fit-for-purpose assessment (Objective 1b).

Objective 1b aimed to assess whether data are fit-for-purpose for conducting safety studies on specific vaccines and outcomes in a near real-time monitoring manner in the future. We assessed the fitness-for-purpose of the data instances using "The Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework" (SPIFD) (27). The SPIFD framework is composed of three steps:

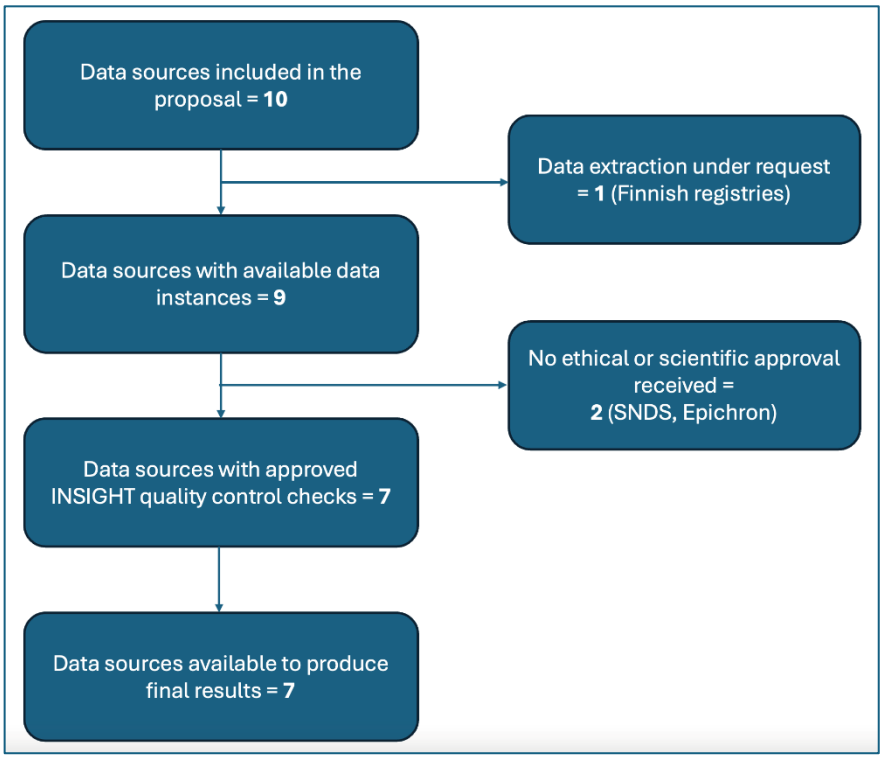
- Operationalization and ranking of minimal criteria variables,
- Identification and narrowing of data sources, and
- Conducting detailed feasibility assessment.

Due to the nature of this work, the criteria to assess the first step is assumed from the study protocol (also presented in the methods section of this report).

10.2.1 IDENTIFICATION AND NARROWING OF DATA SOURCES

Figure 59 depicts the data sources that passed the level 1 and 2 checks and also ran the specific descriptive scripts for disease incidence, prevalence, covariates and vaccine coverage. Ten data sources signed the contract to participate in this study. University of Eastern Finland as the data provider of Finnish data requested access to the required data to the national data holder. This request is under process and therefore, a data instance ready to be included in this report is not yet available. Nine data sources have completed the ETL process to the ConcePTION CDM and passed the INSIGHT level 1 and 2 data quality checks as presented in Table 6 (section 10.1). Two data access providers: Bordeaux Pharmacoepi Platarform (SNDS data) and the Spanish Instituto Aragonés de Ciencias de la Salud (EPICHRON data) have not received the ethical or scientific approval from the corresponding local boards to analyze data specifically for this study. Finally, seven data sources have accomplished the required approvals and requisites to deploy the analysis and produce final results: BIFAP, SIDIAP and VID, from Spain, PEDIANET from Italy, CPRD from the UK, Danish and Norwegian data registries. These data sources have completed the ConcepTION CDM pipeline: ETL'ing a data instance (subset of a data source extracted for the purpose of conducting one or more studies (4)), run and passed the INSIGHT quality checks (see section 10.1.1), and having all the corresponding local committees' approvals. These seven data sources have been assessed in the following step.

Figure 59. Selection of the SAFETY-VAC data sources to be included in this report.



10.2.2 CONDUCTING A DETAILED FEASIBILITY ASSESSMENT

Figure 60 presents a detailed assessment of the availability of required data per data source using the SPIFD tool (27). For the population we required that data were available for the study period (1/1/2017 till most recent) with at least one year look back. Rounding of birth dates means that population cannot be specified well in time and was considered as partial information.

For events, we assumed that availability of in hospital, specialist visit diagnoses and GP diagnoses would render the most complete data (e.g., VID, SIDIAP, NHR). The choice to include primary discharge diagnoses and/or secondary discharge diagnoses (e.g., BIFAP-PC-HOSP), may alter false positive rates and have impacted the absolute rates. For Vaccine exposure we used the WHO/ECDC benchmark data and used the 2019/2020 birth cohorts for assessment.

The assessment also included data access considerations, meaning the time to analyse and produce results using the data instance used in this report or having to ETL and quality check a new data instance within the SAFETY-VAC framework (no need of a new ethical approval). In case of using the current data instance to answer a new research question, all DAPs turned to be fast on conducting, analysing and producing results (max. 3 months). In case the study question requires a new data extraction, BIFAP-ES, SIDIAP-ES, PEDIANET-IT and CPRD-UK are able to produce results in less than 3 months. VID-ES and NHR-NO in 4 to 6 months, and DHR could take more than 6 months.

We cannot make a final fit-for-purpose assessment since this will depend on the specific study question. However, in general, PEDIANET is restricted since it only captures a paediatric population. For several data sources, childhood vaccines were missing in the current data instance; this may change in subsequent extractions. The pregnancy algorithm was not run in all of the data sources and therefore had a low score.

Figure 60. SPIFD heatmap assessment for current data instances

LEGEND		
	5	Many/nearly all data requirements met
	4	Several data requirements met
	3	Likely that several data requirements are met but require further investigation
	2	Some data requirements met or unable to assess at this time
	1	Data requirements not met
	Fast	Fast timelines (e.g., to data access, to analyze) < 3 months
	Moderate	Moderate timelines (e.g., to data access, to analyze) 4 to 6 months
	Slow	Slow timelines (e.g., to data access, to analyze) > 6 months

Study characteristics and considerations	Requested information	Data sources							
		BIFAP_PC-ES	BIFAP_HOSP-ES	SIDIAP-ES	VID-ES	PEDIANET-IT	NHR-NO	DHR-DK	CPRD-UK
DESIGN ELEMENTS									
Study population	• At-least one day of follow-up from 1/1/2017, plus one year look-back. • Age and gender information.	4	4	5	4	5	5	5	5
Vaccine exposure group	Measles-containing vaccines	4	5	5	1	4	5	1	1
	DTP	3	3	5	1	4	5	1	1
	Haemophilus influenzae type B	3	3	5	1	5	5	1	1
	Hepatitis B	3	3	5	1	4	5	1	1
	Polio	3	3	5	1	5	5	1	1
	Pneumococcal conjugate vaccines	5	5	3	1	3	3	1	1
	Varicella	5	5	1	1	5	1	1	1
	HPV	1	1	1	1	4	5	1	1
	Rotavirus	5	5	4	1	4	3	1	1
	Meningoccocal vaccine	5	5	5	1	5	1	1	1
	Influenza vaccine	5	5	5	5	5	5	1	5
	COVID-19 vaccines	5	5	5	5	5	5	5	5
Primary outcomes (availability of events through	Acute coronary artery disease (CAD)	3	3	5	5	1	5	5	4
	ADEM	2	5	5	5	1	5	5	1
	Arrhythmia	5	5	5	5	5	5	5	5
	Arterial thrombosis	3	5	5	5	1	5	5	3

Study characteristics and considerations	Requested information	Data sources							
		BIFAP_PC-ES	BIFAP_HOSP-ES	SIDIAP-ES	VID-ES	PEDIANET-IT	NHR-NO	DHR-DK	CPRD-UK
diagnosis codes and drug proxies)	Autoimmune hepatitis	3	3	3	3	1	3	1	3
	Bell's palsy	5	5	5	5	3	4	5	5
	Cerebral venous sinus thrombosis	2	4	5	5	3	3	3	2
	DIC	1	2	4	4	1	4	4	1
	Erythema multiforme	4	4	5	5	5	5	5	4
	Erythema nodosum	5	5	5	5	5	5	1	1
	Generalized convulsion	3	3	5	5	5	5	5	5
	Haemorrhagic stroke	4	4	5	5	5	5	5	4
	Diabetes type 1	4	4	4	4	5	4	4	4
	Bell's palsy	5	5	5	5	5	5		5
	Grave´s disease	4	4	3	3	4	4	1	1
	Guillain Barré Syndrome	2	5	5	5	5	5	5	2
	Haemorrhagic stroke	4	4	5	5	1	5	5	4
	Hashimoto's thyroiditis	4	4	5	5	4	5	5	4
	Idiopathic thrombocytopenic purpura	4	4	5	5	3	4	4	4
	Kawasaki's disease	4	4	4	4	5	4	4	4
	Meningoencephalitis	2	4	5	5	2	5	4	2
	Microangiopathy	1	3	5	5	2	5	5	1
	Multiple sclerosis	1	4	5	5	1	5	5	1
	Myocarditis	3	4	5	5	1	5	5	4
	Narcolepsy	5	4	5	5	1	5	5	4
	Pancreatitis, acute	3	4	2	2	1	5	5	3
	Pericarditis	4	4	5	5	3	5	5	4
	Polyarteritis nodosa	4	4	5	5	1	5	1	4
	Psoriatic arthropathies	3	3	5	5	1	5	5	1
	Pulmonary embolism	2	4	5	5	1	5	5	4
	Rhabdomyolysis	2	3	4	4	1	3	4	3
	Rheumatoid arthritis	4	4	5	4	3	5	4	4
	SCAR	5	5	5	5	1	5	5	5
	Sensorineural hearing loss	3	3	3	3	3	3	1	3
	Single organ cutaneous vasculitis	4	5	5	5	1	5	5	4
	Systemic lupus erythematosus	4	4	5	4	1	5	5	4
	Thrombocytopenia	3	3	5	5	4	4	3	3

Study characteristics and considerations	Requested information	Data sources							
		BIFAP_PC-ES	BIFAP_HOSP-ES	SIDIAP-ES	VID-ES	PEDIANET-IT	NHR-NO	DHR-DK	CPRD-UK
	Transverse myelitis	3	5	5	5	1	5	5	5
	TTS	2	3	4	4	1	4	5	5
	Ulcerative colitis	5	4	5	5	2	5	5	1
	Venous thromboembolism	4	4	5	5	2	5	5	4
Confounding variables	Availability of key covariates at start of follow-up according to the protocol requirements.	5	5	5	5	5	5	5	5
Key subgroups	Availability to produce a pregnancy cohort	2	2	5	2	1	5	5	2
DATA ACCESS CONSIDERATIONS									
Timeline 1	Time to analyze based on the current instance	Fast	Fast	Fast	Fast	Fast	Fast	Fast	Fast
Timeline 2	Time to analyze based on new data instance within the SAFETY-VAC study framework	Fast	Fast	Fast	Moderate	Fast	Moderate	Slow	Fast

11 DISCUSSION

11.1 Key results

This study was aimed to describe data sources and provide the information that would allow for a fit–for-purpose assessment of the data sources to address potential safety concerns using causal inference. Due to short timelines for providing this quality assessment and fitness for purpose exercise, we have provided data quality checks for 9 data sources which had data instances already available in the ConcePTION CDM; one data access provider was still waiting for data access (Finland). All 9 data sources passed the level 1 (ETL correctness) and level 2 checks (logical checks), while only 4 data sources conducted level 3 checks, mostly because they prioritized to run the dedicated descriptive analytical scripts for this report, which provide a more tailored output for benchmarking of study variables. The analytical scripts were tailored to generate standardized incidence and prevalence rates of 39 events that were predefined by EMA and the study consortium, point prevalence of covariates (risk factors for those events), and vaccine coverage for vaccines included in routine immunization programs in Europe. With exception of some conditions, IRs of events were comparable with prior IRs estimated in ACCESS and/or CVM or published estimates. As we have previously reported in ACCESS (6) and CVM (22), the rates of events differ substantially between data sources with hospital data and GP data versus those with GP data only. This aspect is important to consider when assessing the fit-for-purpose for safety studies of the specific event. The prevalence of covariates was assessed in a 1-year lookback period for diagnosis codes and medicines as proxy. The prevalence of comorbidities was consistent between data sources, although some conditions were not extracted in the current data instances, e.g., for CPRD and DHR. This study also assessed the availability of vaccinations in the data instance and benchmarked with latest WHO and ECDC indicators. Estimation of coverage is a cumulative risk, which is affected by loss to follow-up. Estimating cumulative risk in a dynamic population is challenging due to left and right censoring. We followed two methodological approaches developed in the ADVANCE project, i.e., IPW and PP_{FU}. In that project and after incomplete follow-ups simulated scenarios (creating or not dependence between the period of follow-up and the vaccination), it was recommended to choose a specific method based on the type of censoring and that dependence (11). The SPIFD heatmap shows the fitness for purpose of each data instance in a numerical manner. It is important to realize this is data instance-related and may be changed in a next data instance with more up to date data and an extraction that includes all the required study variables. Since health care data sources change continuously, such renewed ETLs are required. We noticed that in the available data instances, which had mostly been used for COVID-19 research, several childhood vaccines were missing, the population was selected after 2018, and some conditions that were not included in the COVID-19 studies were missing.

11.2 Limitations

This project uses data sources which had available data instances in the ConcePTION CDM at the start of this project (February 15, 2024). Because of the short timelines to deliver the first report (2 months) data instances could not be updated and go through the required quality checks. Data instances can be updated but this requires on average 3 months, as it includes waiting for data to be extracted by the data controller/administrator

and quality checks on the new data instance. We therefore report the available instances which have mostly been used for COVID-19 vaccine studies and assess fitness for purpose specifically for this data instance. Some data sources only included data from 1-1-2019. Although most events were captured, not all childhood vaccines were extracted. Instances contain data for one or more studies and are restricted because of data minimization requirements in the General Data Protection Regulation (GDPR). The GDPR also restricts sharing of small numbers in certain data sources such as Denmark, therefore not all the quality checks can be provided. Data sources include data from different provenances. This has an impact on the ability to identify events, as it depends on the setting where it is diagnosed. This is mentioned in the interpretation of the background rates, and the code counts are provided by type of provenance, to provide full transparency. Comparison with external rates could be challenging since other researchers have not always specified the provenance of the data. Table 5 shows the different types of vocabularies that are utilized by the different DAPs. Harmonization in the ConcePTION pipeline is conducted using script(s), code lists, and the RWE BRIDGE tool (10). All mappings are retrievable and available in the VAC4EU catalogue and study scripts are available in the corresponding GitHub workspaces[5]. Vaccines are captured in different databanks, and codes are often specific to the DAP.

In the IMI-ADVANCE project, a system of mapping was developed, using three letter codes per antigen, this was also used in this project. In this report, we did not produce vaccines coverage information for Mpox, Herpes Zoster and RSV, because these were not ETL'ed in the available data instances. These vaccines will be mapped in next data instances. BIFAP noted that the recorded HPV vaccinations did not include the latest approved 9-valent vaccine (which replaced the previous one, i.e. original Gardasil) and therefore has not been taken into account since it was not mapped into the vxtype (variable into the CDM VACCINE table) (only original Gardasil and Cervarix were mapped) (**Annex 1**). According to external sources, the 9-valent HPV vaccine has been the most used in the last years and has also been administered to males. We also demonstrated that several data sources did not reach the benchmark values, mostly due to registration issues. For fit–for-purpose assessment, we used the 2019 and 2020 birth cohorts since most data sources had recommended end dates at the end of 2022. We used as benchmark the most recent estimates from WHO or ECDC, but these benchmarks may cover different years and include different sources of information. The current report shows the situation of available data instances. To improve the feasibility assessment, we recommend that renewed extractions are done by all DAPs including all vaccines.

Regarding the events, most event rates were comparable with previously reported rates, except for generalized convulsions, thrombocytopenia and SOCV. The code list for the latter had changed based upon clinical discussions following the CVM study. Specifically, BIFAP database includes diagnoses from both primary care consultations recorded through ICPC-2 and ICD-9 by the physicians, and hospital discharge recorded through ICD-9 and ICD-10. ICPC-2 and ICD-9 are internally mapped to SNOMED terms for common data model and analysis purposes. For the current report, 10 events (out of 39) were not completely mapped, and thus not extracted, in the data instance used. Therefore, some patients with the disease recorded only in primary care may not be captured and certain incidence underestimation can be foreseen for Grave´s disease, Hashimoto´s thyroiditis, psoriatic arthropathies, systemic lupus erythematosus, erythema nodosum, auto-immune hepatitis, rheumatoid arthritis, polyarteritis nodosa, ulcerative colitis and multiple sclerosis. Furthermore, code lists for some of these AESI (i.e.,

[5] https://github.com/VAC4EU/ROC18

systemic lupus erythematosus, ulcerative colitis, and rheumatoid arthritis) miss some SNOMED codes (descendants codes of the parent one defining these conditions). Therefore, this may have also contributed to the underestimation of these AESIs' IRs. Next data extraction by BIFAP will include all codes related to these events.
Finally, covariates were available and coherent for all data sources although several medications were not extracted in CPRD, leading to underestimation of comorbidities that had medicines as proxy (e.g., mental health diseases). Moreover, only 3 DAPs ran the pregnancy algorithm and therefore produced a cohort of pregnant persons.

11.3 Interpretation

Seven data sources completed the level 1 and 2 checks and ran the scripts necessary to assess completeness and validity of data on exposure, outcomes and events. This is the status of the current instances. Within the ConcePTION framework, fit-for-purpose assessments were conducted based on the current data instances. We identified several indicators that may be improved, and this can be done with a next ETL process. Therefore, such a fit-for-purpose assessment should be considered per data instance.
Acute and severe AESI that require hospital admission (e.g., DIC or microangiopathy) can possibly be recorded as a secondary hospital discharge diagnosis. Therefore, in such cases, datasources that record only primary diagnosis such as BIFAP_PC_HOSP have lower number of cases. This can be the explanation for the lower incidence rates found for several AESI in BIFAP_PC_HOSP compared with those of other databases from the same country and similar provenances, where all hospital diagnoses were extracted. The decision of discarding secondary diagnosis for future analysis may be reconsidered, especially for AESIs where the sensitivity of secondary diagnoses seems high.

11.4 Generalisability

This study used data from 7 data sources in 5 European countries. Therefore, generalizability of the data is high.

12 OTHER INFORMATION

LIST of ANNEXES

- **Annex 1:** Annex_1_SAFETY-VAC_INSIGHT_Quality_Checks
- **Annex 2:** Annex_2_SAFETY-VAC_Vaccines_Mapping
- **Annex 3:** Annex_3_SAFETY-VAC_Vaccines_Coverage_all_doses
- **Annex 4:** Annex_4_SAFETY-VAC_Standardized_IRs_and_95%CIs
- **Annex 5:** Annex_5_SAFETY-VAC_Age-specific_IRs
- **Annex 6:** Annex_6_SAFETY-VAC_Standardised_Prevalance_Point&Period
- **Annex 7**: Annex_7_SAFETY-VAC_Events_Code_Count_all_Data_Sources

13 CONCLUSIONS

This study has provided a rapid description of the content of 9 data instances from 6 countries. Seven data sources from 5 countries produced final results (attrition, prevalence of covariates, coverage of selected vaccines, and incidence and prevalence rates of 39 outcomes). These data instances did not always contain the variables that were required for the study, as they had been initially prepared for other projects, i.e. ConcePTION or COVID-19 studies. This will be updated in new data extractions. This fit-for-purpose assessment was performed per database by applying the SPIFD tool(27). Three data sources could not provide the required data in time for this report due to slow ethical or scientific approvals (SNDS, EPICHRON), and lengthy waiting times to retrieve data (Finland). To be ready to address new safety questions for the VMP, while respecting the fact that DAPs may not request all of the data from controllers due to GDPR minimization criteria, it would be recommendable to limit in calendar time, or create specific cohorts that would be covering vaccination target groups, and ETL all events, all medicines, all vaccine records, and all medical/survey observations.

14 REFERENCES

1. Salmon DA, Lambert PH, Nohynek HM, Gee J, Parashar UD, Tate JE, et al. Novel vaccine safety issues and areas that would benefit from further research. BMJ Glob Health [Internet]. 2021 May 1 [cited 2024 May 8];6(Suppl 2):e003814. Available from: https://gh.bmj.com/content/6/Suppl_2/e003814
2. Ricotta EE, Rid A, Cohen IG, Evans NG. Observational studies must be reformed before the next pandemic. Nature Medicine 2023 29:8 [Internet]. 2023 Jun 7 [cited 2024 May 8];29(8):1903–5. Available from: https://www.nature.com/articles/s41591-023-02375-8
3. Vaccine Monitoring Platform | European Medicines Agency [Internet]. [cited 2024 May 8]. Available from: https://www.ema.europa.eu/en/about-us/what-we-do/crisis-preparedness-management/vaccine-monitoring-platform
4. Thurin NH, Pajouheshnia R, Roberto G, Dodd C, Hyeraci G, Bartolini C, et al. From Inception to ConcePTION: Genesis of a Network to Support Better Monitoring and Communication of Medication Safety During Pregnancy and Breastfeeding. Clin Pharmacol Ther [Internet]. 2022 Jan 1 [cited 2024 May 8];111(1):321–31. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/cpt.2476
5. Becker BF, Kors JA, van Mulligen EM, Sturkenboom MC. Alignment of vaccine codes using an ontology of vaccine descriptions. J Biomed Semantics [Internet]. 2022 Dec 1 [cited 2024 May 8];13(1):1–12. Available from: https://jbiomedsem.biomedcentral.com/articles/10.1186/s13326-022-00278-0
6. Willame C, Dodd C, Durán CE, Elbers RJHJ, Gini R, Bartolini C, et al. Background rates of 41 adverse events of special interest for COVID-19 vaccines in 10 European healthcare databases - an ACCESS cohort study. Vaccine. 2023 Jan 4;41(1):251–62.
7. Zenodo [Internet]. [cited 2024 May 8]. Available from: https://zenodo.org/
8. Becker BFH, Avillach P, Romio S, van Mulligen EM, Weibel D, Sturkenboom MCJM, et al. CodeMapper: semiautomatic coding of case definitions. A contribution from the ADVANCE project. Pharmacoepidemiol Drug Saf [Internet]. 2017 Aug 1 [cited 2024 May 23];26(8):998–1005. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/pds.4245
9. ARS-toscana/ConcePTIONAlgorithmPregnancies: Repository of the script of the ConcePTION Algorithm for Pregnancies [Internet]. [cited 2024 May 10]. Available from: https://github.com/ARS-toscana/ConcePTIONAlgorithmPregnancies
10. Royo AC, Elbers R, Weibel D, Hoxhaj V, Kurkcuoglu Z, Sturkenboom MC, et al. Real-World Evidence BRIDGE: a tool to connect protocol with code programming. medRxiv [Internet]. 2024 May 8 [cited 2024 May 10];2024.05.08.24306833. Available from: https://www.medrxiv.org/content/10.1101/2024.05.08.24306833v1
11. Braeye T, Bauchau V, Sturkenboom M, Emborg HD, García AL, Huerta C, et al. Estimation of vaccination coverage from electronic healthcare records; methods performance evaluation – A contribution of the ADVANCE-project. PLoS One [Internet]. 2019 Sep 1 [cited 2024 May 10];14(9):e0222296. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0222296
12. Braeye T, Emborg HD, Llorente-García A, Huerta C, Martín-Merino E, Duarte-Salles T, et al. Age-specific vaccination coverage estimates for influenza, human papillomavirus and measles containing vaccines from seven population-based healthcare databases from four EU countries – The ADVANCE project. Vaccine. 2020 Apr 3;38(16):3243–54.
13. Measles-containing-vaccine first-dose (MCV1) immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/measles-containing-vaccine-first-dose-(mcv1)-immunization-coverage-among-1-year-olds-(-)
14. Diphtheria tetanus toxoid and pertussis (DTP3) immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/diphtheria-tetanus-toxoid-and-pertussis-(dtp3)-immunization-coverage-among-1-year-olds-(-)
15. Hib (Hib3) immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hib-hib3-immunization-coverage-among-1-year-olds
16. Hepatitis B (HepB3) immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hepatitis-b-(hepb3)-immunization-coverage-among-1-year-olds-(-)

17. Polio (Pol3) immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/polio-(pol3)-immunization-coverage-among-1-year-olds-(-)
18. Pneumococcal conjugate vaccines (PCV3) immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/pneumoccocal-conjugate-vaccines-(pcv3)-immunization-coverage-among-1-year-olds-(-)
19. HPV immunization coverage estimates among primary target cohort (9-14 years old girls) (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/girls-aged-15-years-old-that-received-the-recommended-doses-of-hpv-vaccine
20. Rotavirus vaccines completed dose (RotaC) immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/rotavirus-vaccines-completed-dose-(rotac)-immunization-coverage-among-1-year-olds-(-)
21. BCG immunization coverage among 1-year-olds (%) [Internet]. [cited 2024 May 8]. Available from: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/bcg-immunization-coverage-among-1-year-olds-(-)
22. Durán CE, Messina D, Gini R, Riefolo F, Aragón M, Belitser S, et al. Rapid Safety Assessment of SARS-CoV-2 Vaccines in EU Member States using Electronic Health Care Data Sources (COVID Vaccine Monitor-CVM study): Final Study Report for WP3 (electronic health record data). 2023 Aug 24;
23. Hoxhaj V, Navarro CLA, Riera-Arnau J, Elbers RJ, Alsina E, Dodd C, et al. INSIGHT: A Tool for Fit-for-Purpose Evaluation and Quality Assessment of Observational Data Sources for Real World Evidence on Medicine and Vaccine Safety. medRxiv [Internet]. 2023 Oct 30 [cited 2024 May 8];2023.10.30.23297753. Available from: https://www.medrxiv.org/content/10.1101/2023.10.30.23297753v1
24. UMC-Utrecht-RWE/INSIGHT-Level1: V1. [cited 2024 May 8]; Available from: https://zenodo.org/records/10035167
25. UMC-Utrecht-RWE/INSIGHT-Level2: V1. [cited 2024 May 8]; Available from: https://zenodo.org/records/10035169
26. UMC-Utrecht-RWE/INSIGHT-Level3: V1. [cited 2024 May 8]; Available from: https://zenodo.org/records/10035171
27. Gatto NM, Campbell UB, Rubinstein E, Jaksa A, Mattox P, Mo J, et al. The Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework. Clin Pharmacol Ther [Internet]. 2022 Jan 1 [cited 2024 May 8];111(1):122–34. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/cpt.2466
28. Data Quality Framework for EU medicines regulation. [cited 2024 May 8]; Available from: www.ema.europa.eu/contact
29. Global Health Observatory [Internet]. [cited 2024 May 22]. Available from: https://www.who.int/data/gho
30. COVID-19 Vaccine Tracker | European Centre for Disease Prevention and Control [Internet]. [cited 2024 May 22]. Available from: https://vaccinetracker.ecdc.europa.eu/public/extensions/COVID-19/vaccine-tracker.html#summary-tab
31. Pozza F, Piovesan C, Russo F, Bella A, Pezzotti P, Emberti Gialloreti L. Impact of universal vaccination on the epidemiology of varicella in Veneto, Italy. Vaccine. 2011 Nov 28;29(51):9480–7.
32. Tafuri S, Fortunato F, Cappelli MG, Cozza V, Bechini A, Bonanni P, et al. Effectiveness of vaccination against varicella in children under 5 years in Puglia, Italy 2006–2012. Hum Vaccin Immunother [Internet]. 2015 Jan 1 [cited 2024 May 8];11(1):214–9. Available from: https://www.tandfonline.com/doi/abs/10.4161/hv.36153
33. Consulta Interactiva del SNS [Internet]. [cited 2024 May 22]. Available from: https://pestadistico.inteligenciadegestion.sanidad.gob.es/publicoSNS/I/sivamin/sivamin
34. BCG [Internet]. [cited 2024 May 27]. Available from: https://www.who.int/teams/health-product-policy-and-standards/standards-and-specifications/vaccines-quality/bcg
35. Emborg HD, Kahlert J, Braeye T, Bauwens J, Bollaerts K, Danieli G, et al. ADVANCE system testing: Can coverage of pertussis vaccination be estimated in European countries using electronic healthcare databases: An example. Vaccine. 2020 Dec 22;38:B22–30.

36. Ruiz-Contreras J, Alfayate-Miguelez S, Carazo-Gallego B, Onís E, Díaz-Munilla L, Mendizabal M, et al. Rotavirus gastroenteritis hospitalizations in provinces with different vaccination coverage rates in Spain, 2013–2018. BMC Infect Dis [Internet]. 2021 Dec 1 [cited 2024 May 8];21(1):1–7. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-021-06841-x
37. Meningococcal vaccination coverage [Internet]. [cited 2024 May 8]. Available from: https://immunizationdata.who.int/global/wiise-detail-page/meningococcal-vaccination-coverage?CODE=GBR&ANTIGEN=&YEAR=
38. Seasonal influenza vaccination recommendations and coverage rates in EU/EEA Member States. [cited 2024 May 8]; Available from: https://www.ecdc.europa.eu/en/publications-data/seasonal-influenza-vaccination-recommendations-and-coverage-rates-eueea-member
39. First COVID-19 vaccine approved for children aged 12 to 15 in EU | European Medicines Agency [Internet]. [cited 2024 May 29]. Available from: https://www.ema.europa.eu/en/news/first-covid-19-vaccine-approved-children-aged-12-15-eu
40. COVID-19 vaccine Spikevax approved for children aged 12 to 17 in EU | European Medicines Agency [Internet]. [cited 2024 May 29]. Available from: https://www.ema.europa.eu/en/news/covid-19-vaccine-spikevax-approved-children-aged-12-17-eu
41. SIVIC - COVID-19 - Vacunació [Internet]. [cited 2024 May 29]. Available from: https://sivic.salut.gencat.cat/covid_vacunacio
42. Coronary Artery Disease - MeSH - NCBI [Internet]. [cited 2024 May 8]. Available from: https://www.ncbi.nlm.nih.gov/mesh/68003324
43. Kelters I, Willame C, Durán C, Belbachir L, Martín-Pérez M, García-Poza P, et al. ACCESS-Background rate of adverse events-definition –Coronary Artery Disease. 2021 [cited 2024 May 8]; Available from: https://zenodo.org/records/5226602
44. Sundaram V, Bloom C, Zakeri R, Halcox J, Cohen A, Bowrin K, et al. Temporal trends in the incidence, treatment patterns, and outcomes of coronary artery disease and peripheral artery disease in the UK, 2006–2015. Eur Heart J [Internet]. 2020 May 1 [cited 2024 May 28];41(17):1636–49. Available from: https://dx.doi.org/10.1093/eurheartj/ehz880
45. Encephalomyelitis, Acute Disseminated - MeSH - NCBI [Internet]. [cited 2024 May 8]. Available from: https://www.ncbi.nlm.nih.gov/mesh/?term=acute+disseminated+encephalomyelitis
46. ACCESS: Background rates of AESI to monitor vaccine safety- definition Acute disseminated encephalomyelitis. [cited 2024 May 22]; Available from: https://zenodo.org/records/5109555
47. Morgan HJ, Clothier HJ, Sepulveda Kattan G, Boyd JH, Buttery JP. Acute disseminated encephalomyelitis and transverse myelitis following COVID-19 vaccination – A self-controlled case series analysis. Vaccine. 2024 Apr 2;42(9):2212–9.
48. Gubernot D, Jazwa A, Niu M, Baumblatt J, Gee J, Moro P, et al. U.S. Population-Based background incidence rates of medical conditions for use in safety assessment of COVID-19 vaccines. Vaccine. 2021 Jun 23;39(28):3666–77.
49. Pavone P, Pettoello-Mantovano M, Le Pira A, Giardino I, Pulvirenti A, Giugno R, et al. Acute disseminated encephalomyelitis: A long-term prospective study and meta-analysis. Neuropediatrics [Internet]. 2010 [cited 2024 May 22];41(6):246–55. Available from: http://www.thieme-connect.de/products/ejournals/html/10.1055/s-0031-1271656
50. Willame C, Dodd C, van der Aa L, Picelli G, Emborg HD, Kahlert J, et al. Incidence Rates of Autoimmune Diseases in European Healthcare Databases: A Contribution of the ADVANCE Project. Drug Saf [Internet]. 2021 Mar 1 [cited 2024 May 22];44(3):383–95. Available from: https://link.springer.com/article/10.1007/s40264-020-01031-1
51. ACCESS-Background rate of adverse events-definition –Arrhythmia. [cited 2024 May 22]; Available from: https://zenodo.org/records/5226644
52. Martinez C, Katholing A, Wallenhorst C, Granziera S, Cohen AT, Freedman SB. Increasing incidence of non-valvular atrial fibrillation in the UK from 2001 to 2013. Heart [Internet]. 2015 Nov 1 [cited 2024 May 26];101(21):1748–54. Available from: https://heart.bmj.com/content/101/21/1748
53. Linz D, Gawalko M, Betz K, Hendriks JM, Lip GYH, Vinter N, et al. Atrial fibrillation: epidemiology, screening and digital health. The Lancet Regional Health - Europe [Internet]. 2024 Feb 1 [cited 2024 May 26];37:100786. Available from: http://www.thelancet.com/article/S2666776223002053/fulltext
54. Hegelund ER, Kjerpeseth LJ, Mortensen LH, Igland J, Berge T, Anjum M, et al. Prevalence and Incidence Rates of Atrial Fibrillation in Denmark 2004–2018. Clin Epidemiol [Internet]. 2022 Oct 26 [cited 2024 May 26];14:1193–204. Available from:

https://www.dovepress.com/prevalence-and-incidence-rates-of-atrial-fibrillation-in-denmark-20042-peer-reviewed-fulltext-article-CLEP
55. Choi J, Kermode JC. New therapeutic approaches to combat arterial thrombosis. Mol Interv. 2011 Apr;11(2):111–23.
56. Wendelboe AM, Raskob GE. Global Burden of Thrombosis. Circ Res [Internet]. 2016 Apr 29 [cited 2024 May 26];118(9):1340–7. Available from: https://www.ahajournals.org/doi/abs/10.1161/CIRCRESAHA.115.306841
57. Mundal LJ, Hovland A, Igland J, Vetrhus M, Veierød MB, Holven KB, et al. Increased risk of peripheral artery disease in persons with familial hypercholesterolaemia: a prospective registry study. Eur J Prev Cardiol [Internet]. 2022 Feb 9 [cited 2024 May 26];28(18):e11–3. Available from: https://dx.doi.org/10.1093/eurjpc/zwaa024
58. Søgaard M, Nielsen PB, Eldrup N, Behrendt CA, Nicolajsen CW, Lip GYH, et al. Epidemiological Trends and Projections of Incidence, Prevalence, and Disease Related Mortality Associated With Peripheral Arterial Disease: Observations Using Nationwide Danish Data. European Journal of Vascular and Endovascular Surgery [Internet]. 2023 Nov 1 [cited 2024 May 26];66(5):662–9. Available from: http://www.ejves.com/article/S1078588423006287/fulltext
59. Czaja AJ. Frequency and nature of the variant syndromes of autoimmune liver disease. Hepatology [Internet]. 1998 Aug 1 [cited 2024 May 8];28(2):360–5. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/hep.510280210
60. Hahn JW, Yang HR, Moon JS, Chang JY, Lee K, Kim GA, et al. Global incidence and prevalence of autoimmune hepatitis, 1970–2022: a systematic review and meta-analysis. EClinicalMedicine [Internet]. 2023 Nov 1 [cited 2024 May 26];65:102280. Available from: http://www.thelancet.com/article/S2589537023004571/fulltext
61. [Incidence and prevalence of autoimmune liver diseases] - PubMed [Internet]. [cited 2024 May 26]. Available from: https://pubmed.ncbi.nlm.nih.gov/9889636/
62. Grønbæk L, Vilstrup H, Jepsen P. Autoimmune hepatitis in Denmark: Incidence, prevalence, prognosis, and causes of death. A nationwide registry-based cohort study. J Hepatol [Internet]. 2014 Mar 1 [cited 2024 May 26];60(3):612–7. Available from: http://www.journal-of-hepatology.eu/article/S0168827813007459/fulltext
63. Lv T, Li M, Zeng N, Zhang J, Li S, Chen S, et al. Systematic review and meta-analysis on the incidence and prevalence of autoimmune hepatitis in Asian, European, and American population. J Gastroenterol Hepatol [Internet]. 2019 Oct 1 [cited 2024 May 22];34(10):1676–84. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/jgh.14746
64. Kochhar S, Assis DN, Mack C, Izurieta HS, Muratori L, Munoz A, et al. Autoimmune hepatitis: Brighton Collaboration case definition and guidelines for data collection, analysis, and presentation of immunisation safety data. Vaccine. 2024 Mar 7;42(7):1812–25.
65. Finsterer J. Management of peripheral facial nerve palsy. European Archives of Oto-Rhino-Laryngology [Internet]. 2008 Jul 27 [cited 2024 May 26];265(7):743–52. Available from: https://link.springer.com/article/10.1007/s00405-008-0646-4
66. Monini S, Lazzarino AI, Iacolucci C, Buffoni A, Barbara M. Epidemiology of Bell's palsy in an Italian Health District: incidence and case-control study. Acta Otorhinolaryngologica Italica [Internet]. 2010 Aug [cited 2024 May 26];30(4):198–204. Available from: /pmc/articles/PMC3008145/
67. Kim MH, Park SY. Population-based study and a scoping review for the epidemiology and seasonality in and effect of weather on Bell's palsy. Scientific Reports 2021 11:1 [Internet]. 2021 Aug 20 [cited 2024 May 26];11(1):1–9. Available from: https://www.nature.com/articles/s41598-021-96422-4
68. Li X, Ostropolets A, Makadia R, Shoaibi A, Rao G, Sena AG, et al. Characterising the background incidence rates of adverse events of special interest for covid-19 vaccines in eight countries: multinational network cohort study. BMJ [Internet]. 2021 Jun 14 [cited 2024 May 22];373. Available from: https://www.bmj.com/content/373/bmj.n1435
69. Li X, Ostropolets A, Makadia R, Shoaibi A, Rao G, Sena AG, et al. Characterising the background incidence rates of adverse events of special interest for covid-19 vaccines in eight countries: multinational network cohort study. BMJ [Internet]. 2021 Jun 14 [cited 2024 May 28];373. Available from: https://www.bmj.com/content/373/bmj.n1435
70. Phillips A, Jiang Y, Walsh D, Andrews N, Artama M, Clothier H, et al. Background rates of adverse events of special interest for COVID-19 vaccines: A multinational Global Vaccine Data Network (GVDN) analysis. Vaccine. 2023 Oct 6;41(42):6227–38.
71. Meng R, Dornbos D, Meng L, Wu Y, Liu Y, Li G, et al. Clinical differences between acute CVST and non-thrombotic CVSS. Clin Neurol Neurosurg. 2012 Nov 1;114(9):1257–62.

72. Pottegård A, Lund LC, Karlstad Ø, Dahl J, Andersen M, Hallas J, et al. Arterial events, venous thromboembolism, thrombocytopenia, and bleeding after vaccination with Oxford-AstraZeneca ChAdOx1-S in Denmark and Norway: population based cohort study. BMJ [Internet]. 2021 May 5 [cited 2024 May 10];373. Available from: https://www.bmj.com/content/373/bmj.n1114
73. Pottegård A, Lund LC, Karlstad Ø, Dahl J, Andersen M, Hallas J, et al. Arterial events, venous thromboembolism, thrombocytopenia, and bleeding after vaccination with Oxford-AstraZeneca ChAdOx1-S in Denmark and Norway: population based cohort study. BMJ [Internet]. 2021 May 5 [cited 2024 May 28];373. Available from: https://www.bmj.com/content/373/bmj.n1114
74. Association AD. Diagnosis and Classification of Diabetes Mellitus. Diabetes Care [Internet]. 2009 Jan 1 [cited 2024 May 8];32(Supplement_1):S62–7. Available from: https://dx.doi.org/10.2337/dc09-S062
75. Mobasseri M, Shirmohammadi M, Amiri T, Vahed N, Fard HH, Ghojazadeh M. Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis. Health Promot Perspect [Internet]. 2020 [cited 2024 May 28];10(2):98. Available from: /pmc/articles/PMC7146037/
76. Rasmussen TA, Jørgensen MRS, Bjerrum S, Jensen-Fangel S, Støvring H, Østergaard L, et al. Use of population based background rates of disease to assess vaccine safety in childhood and mass immunisation in Denmark: nationwide population based cohort study. The BMJ [Internet]. 2012 Oct 13 [cited 2024 May 28];345(7878). Available from: /pmc/articles/PMC3444137/
77. Carstensen B, Rønn PF, Jørgensen ME. Prevalence, incidence and mortality of type 1 and type 2 diabetes in Denmark 1996–2016. BMJ Open Diabetes Res Care [Internet]. 2020 May 1 [cited 2024 May 28];8(1):1071. Available from: /pmc/articles/PMC7265004/
78. Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of Type 1 Diabetes. Endocrinol Metab Clin North Am. 2010 Sep 1;39(3):481–97.
79. Boral BM, Williams DJ, Boral LI. Disseminated Intravascular Coagulation. Am J Clin Pathol [Internet]. 2016 Dec 1 [cited 2024 May 9];146(6):670–80. Available from: https://dx.doi.org/10.1093/ajcp/aqw195
80. Cullen LA, Grange Z, Antal K, Waugh L, Alsina MS, Gibbons CL, et al. COVID-19 vaccine safety in Scotland – background rates of adverse events of special interest. Public Health. 2023 Nov 1;224:1–7.
81. Erythema Multiforme - MeSH - NCBI [Internet]. [cited 2024 May 22]. Available from: https://www.ncbi.nlm.nih.gov/mesh/?term=erythema+multiforme
82. Samim F, Auluck A, Zed C, Williams PM. Erythema Multiforme: A Review of Epidemiology, Pathogenesis, Clinical Features, and Treatment. Dent Clin North Am. 2013 Oct 1;57(4):583–96.
83. ACCESS-Background rate of adverse events-definition –Erythema Multiforma. [cited 2024 May 22]; Available from: https://zenodo.org/records/5236231
84. Scully C. Erythema multiforme. Oral and Maxillofacial Medicine [Internet]. 2013 Jan 1 [cited 2024 May 26];271–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780702049484000428
85. Pérez-Garza DM, Chavez-Alvarez S, Ocampo-Candiani J, Gomez-Flores M. Erythema Nodosum: A Practical Approach and Diagnostic Algorithm. Am J Clin Dermatol [Internet]. 2021 May 1 [cited 2024 May 22];22(3):367–78. Available from: https://link.springer.com/article/10.1007/s40257-021-00592-w
86. Blake T, Manahan M, Rodins K. Erythema nodosum - A review of an uncommon panniculitis. Dermatol Online J [Internet]. 2014 [cited 2024 May 22];20(4). Available from: https://escholarship.org/uc/item/2dt0z3mz
87. Leung AKC, Leong KF, Lam JM. Erythema nodosum. World Journal of Pediatrics [Internet]. 2018 Dec 1 [cited 2024 May 10];14(6):548–54. Available from: https://link.springer.com/article/10.1007/s12519-018-0191-1
88. Clinical, Epidemiological, and Etiological Changes in Erythema Nodosum - PubMed [Internet]. [cited 2024 May 22]. Available from: https://pubmed.ncbi.nlm.nih.gov/30550008/
89. Erythema nodosum in the Negev area--a survey of 50 patients - PubMed [Internet]. [cited 2024 May 22]. Available from: https://pubmed.ncbi.nlm.nih.gov/3440746/
90. Beghi E. The Epidemiology of Epilepsy. Neuroepidemiology [Internet]. 2020 Mar 17 [cited 2024 May 26];54(2):185–91. Available from: https://dx.doi.org/10.1159/000503831
91. Martinez C, Sullivan T, Hauser WA. Prevalence of acute repetitive seizures (ARS) in the United Kingdom. Epilepsy Res [Internet]. 2009 Dec [cited 2024 May 10];87(2–3):137–43. Available from: https://pubmed.ncbi.nlm.nih.gov/19748226/
92. Christensen J, Vestergaard M, Pedersen MG, Pedersen CB, Olsen J, Sidenius P. Incidence and prevalence of epilepsy in Denmark. Epilepsy Res. 2007 Aug 1;76(1):60–5.

93. Chaker L, Cooper DS, Walsh JP, Peeters RP. Hyperthyroidism. The Lancet [Internet]. 2024 Feb 24 [cited 2024 May 14];403(10428):768–80. Available from: http://www.thelancet.com/article/S0140673623020160/fulltext
94. Chaker L, Cooper DS, Walsh JP, Peeters RP. Hyperthyroidism. The Lancet [Internet]. 2024 Feb 24 [cited 2024 May 28];403(10428):768–80. Available from: http://www.thelancet.com/article/S0140673623020160/fulltext
95. Sophia EA, Folkestad L, Brix T, Joop BS. Long-Term outcome of the thyroid function in graves. Endocrine Abstracts [Internet]. 2023 Aug 24 [cited 2024 May 28];92. Available from: https://www.endocrine-abstracts.org/ea/0092/ea0092op-04-03
96. Metwalley KA, Farghaly HS. Graves' Disease in Children: An Update. Clin Med Insights Endocrinol Diabetes [Internet]. 2023 Jan 1 [cited 2024 May 29];16. Available from: /pmc/articles/PMC10161304/
97. ACCESS: Background rates of AESI to monitor vaccine safety- GBS definition. [cited 2024 May 10]; Available from: https://zenodo.org/records/5109436
98. Bragazzi NL, Kolahi AA, Nejadghaderi SA, Lochner P, Brigo F, Naldi A, et al. Global, regional, and national burden of Guillain–Barré syndrome and its underlying causes from 1990 to 2019. J Neuroinflammation [Internet]. 2021 Dec 1 [cited 2024 May 29];18(1):1–11. Available from: https://jneuroinflammation.biomedcentral.com/articles/10.1186/s12974-021-02319-4
99. Govoni V, Granieri E. Epidemiology of the Guillain-Barré syndrome. Curr Opin Neurol [Internet]. 2001 [cited 2024 May 10];14(5):605–13. Available from: https://pubmed.ncbi.nlm.nih.gov/11562572/
100. Levison LS, Thomsen RW, Christensen DH, Mellemkjær T, Sindrup SH, Andersen H. Guillain-Barré syndrome in Denmark: validation of diagnostic codes and a population-based nationwide study of the incidence in a 30-year period. Clin Epidemiol [Internet]. 2019 [cited 2024 May 29];11:275–83. Available from: https://pubmed.ncbi.nlm.nih.gov/31114387/
101. Rasmussen TA, Jørgensen MRS, Bjerrum S, Jensen-Fangel S, Støvring H, Østergaard L, et al. Use of population based background rates of disease to assess vaccine safety in childhood and mass immunisation in Denmark: nationwide population based cohort study. BMJ [Internet]. 2012 Oct 13 [cited 2024 May 29];345(7878). Available from: https://pubmed.ncbi.nlm.nih.gov/22988304/
102. Farbu E, Rudolph T, Stefansdottir S. Guillain Barré syndrome--Incidence and clinical course in Southern Rogaland, Norway. Clin Neurol Neurosurg [Internet]. 2016 Feb 1 [cited 2024 May 29];141:33–7. Available from: https://pubmed.ncbi.nlm.nih.gov/26724429/
103. Carey IM, Banchoff E, Nirmalananthan N, Harris T, DeWilde S, Chaudhry UAR, et al. Prevalence and incidence of neuromuscular conditions in the UK between 2000 and 2019: A retrospective study using primary care data. PLoS One [Internet]. 2021 Dec 1 [cited 2024 May 29];16(12). Available from: https://pubmed.ncbi.nlm.nih.gov/34972157/
104. Carey IM, Banchoff E, Nirmalananthan N, Harris T, DeWilde S, Chaudhry UAR, et al. Prevalence and incidence of neuromuscular conditions in the UK between 2000 and 2019: A retrospective study using primary care data. PLoS One [Internet]. 2021 Dec 1 [cited 2024 May 29];16(12). Available from: https://pubmed.ncbi.nlm.nih.gov/34972157/
105. AESI Case Definition Companion Guide: Guillain Barré and Miller Fisher Syndrome. [cited 2024 May 28]; Available from: https://zenodo.org/records/6668879
106. Grysiewicz RA, Thomas K, Pandey DK. Epidemiology of Ischemic and Hemorrhagic Stroke: Incidence, Prevalence, Mortality, and Risk Factors. Neurol Clin. 2008 Nov 1;26(4):871–95.
107. Purroy F, Montalà N. Epidemiology of stroke in the last decade: a systematic review. Rev Neurol. 2021 Nov 1;73(9):321–36.
108. McLeod DSA, Cooper DS. The incidence and prevalence of thyroid autoimmunity. Endocrine [Internet]. 2012 [cited 2024 May 10];42(2):252–65. Available from: https://pubmed.ncbi.nlm.nih.gov/22644837/
109. Galofré JC, García-Mayor R V., Fluiters E, Fernàndez-Calvet L, Rego A, Pàramo C, et al. Incidence of different forms of thyroid dysfunction and its degrees in an iodine sufficient area. Thyroidology [Internet]. 1994 Aug 1 [cited 2024 May 27];6(2):49–54. Available from: https://europepmc.org/article/med/7536450
110. Flynn RWV, MacDonald TM, Morris AD, Jung RT, Leese GP. The Thyroid Epidemiology, Audit, and Research Study: Thyroid Dysfunction in the General Population. J Clin Endocrinol Metab [Internet]. 2004 Aug 1 [cited 2024 May 27];89(8):3879–84. Available from: https://dx.doi.org/10.1210/jc.2003-032089
111. Cerqueira C, Knudsen N, Ovesen L, Laurberg P, Perrild H, Rasmussen LB, et al. Doubling in the use of thyroid hormone replacement therapy in Denmark: Association to iodization of salt? Eur J

Epidemiol [Internet]. 2011 Aug 10 [cited 2024 May 29];26(8):629–35. Available from: https://link.springer.com/article/10.1007/s10654-011-9590-5
112. Carlé A, Laurberg P, Pedersen IB, Knudsen N, Perrild H, Ovesen L, et al. Epidemiology of subtypes of hypothyroidism in Denmark. Eur J Endocrinol [Internet]. 2006 Jan 1 [cited 2024 May 29];154(1):21–8. Available from: https://dx.doi.org/10.1530/eje.1.02068
113. ACCESS-Background rate of adverse events-definition -thrombocytopenia. [cited 2024 May 22]; Available from: https://zenodo.org/records/5169150
114. Christiansen CF, Bahmanyar S, Ghanima W, Risbo N, Ekstrand C, Stryker S, et al. Chronic immune thrombocytopenia in Denmark, Sweden and Norway: The Nordic Country Patient Registry for Romiplostim. EClinicalMedicine [Internet]. 2019 Sep 1 [cited 2024 May 27];14:80–7. Available from: http://www.thelancet.com/article/S258953701930135X/fulltext
115. Nasreen S, Calzavara A, Buchan SA, Thampi N, Johnson C, Wilson SE, et al. Background incidence rates of adverse events of special interest related to COVID-19 vaccines in Ontario, Canada, 2015 to 2020, to inform COVID-19 vaccine safety surveillance. Vaccine. 2022 May 26;40(24):3305–12.
116. Gubernot D, Jazwa A, Niu M, Baumblatt J, Gee J, Moro P, et al. U.S. Population-Based background incidence rates of medical conditions for use in safety assessment of COVID-19 vaccines. Vaccine. 2021 Jun 23;39(28):3666–77.
117. Phuong LK, Bonetto C, Buttery J, Pernus YB, Chandler R, Goldenthal KL, et al. Kawasaki disease and immunisation: Standardised case definition & guidelines for data collection, analysis. Vaccine [Internet]. 2016 Dec 12 [cited 2024 May 10];34(51):6582–96. Available from: https://pubmed.ncbi.nlm.nih.gov/27863715/
118. Piram M. Epidemiology of Kawasaki Disease in Europe. Front Pediatr [Internet]. 2021 May 25 [cited 2024 May 10];9. Available from: https://pubmed.ncbi.nlm.nih.gov/34113590/
119. Sejvar JJ, Kohl KS, Bilynsky R, Blumberg D, Cvetkovich T, Galama J, et al. Encephalitis, myelitis, and acute disseminated encephalomyelitis (ADEM): case definitions and guidelines for collection, analysis, and presentation of immunization safety data. Vaccine [Internet]. 2007 Aug 1 [cited 2024 May 10];25(31):5771–92. Available from: https://pubmed.ncbi.nlm.nih.gov/17570566/
120. George BP, Schneider EB, Venkatesan A. Encephalitis Hospitalization Rates and Inpatient Mortality in the United States, 2000-2010. PLoS One [Internet]. 2014 Sep 1 [cited 2024 May 27];9(9):e104169. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104169
121. ACCESS-Background rate of adverse events-definition –Microangiopathy. [cited 2024 May 22]; Available from: https://zenodo.org/records/5169451
122. Belbasis L, Bellou V, Evangelou E, Ioannidis JPA, Tzoulaki I. Environmental risk factors and multiple sclerosis: An umbrella review of systematic reviews and meta-analyses. Lancet Neurol [Internet]. 2015 Mar 1 [cited 2024 May 15];14(3):263–73. Available from: http://www.thelancet.com/article/S1474442214702674/fulltext
123. Adams RD, Victor M, Ropper AH. PRINCIPLES OF NEUROLOGY. 6TH EDITION. 1997. By Raymond D. Adams, Maurice Victor, Allan H. Ropper. Published by McGraw Hill. 1618 pages. Canadian Journal of Neurological Sciences [Internet]. 1997 Nov [cited 2024 May 15];24(4):363–363. Available from: https://www.cambridge.org/core/journals/canadian-journal-of-neurological-sciences/article/principles-of-neurology-6th-edition-1997-by-raymond-d-adams-maurice-victor-allan-h-ropper-published-by-mcgraw-hill-1618-pages/5A9C1D1DC386B1429C912107F977951F
124. Pugliatti M, Rosati G, Carton H, Riise T, Drulovic J, Vécsei L, et al. The epidemiology of multiple sclerosis in Europe. Eur J Neurol [Internet]. 2006 Jul 1 [cited 2024 May 10];13(7):700–22. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1468-1331.2006.01342.x
125. Alcalde-Cabero E, Almazán-Isla J, García-Merino A, de Sá J, de Pedro-Cuesta J. Incidence of multiple sclerosis among European Economic Area populations, 1985-2009: The framework for monitoring. BMC Neurol [Internet]. 2013 Jun 12 [cited 2024 May 10];13(1):1–22. Available from: https://bmcneurol.biomedcentral.com/articles/10.1186/1471-2377-13-58
126. Magyari M, Joensen H, Laursen B, Koch-Henriksen N. The Danish Multiple Sclerosis Registry. Brain Behav [Internet]. 2021 Jan 1 [cited 2024 May 27];11(1):e01921. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/brb3.1921
127. Lane J, Ng HS, Poyser C, Lucas RM, Tremlett H. Multiple sclerosis incidence: A systematic review of change over time by geographical region. Mult Scler Relat Disord [Internet]. 2022 Jul 1 [cited 2024 May 10];63. Available from: http://www.msard-journal.com/article/S2211034822004436/fulltext

128. ACCESS-Background rate of adverse events-definition –Myocarditis and/or pericarditis. [cited 2024 May 22]; Available from: https://zenodo.org/records/5172798
129. Costa C, Moniati F. The Epidemiology of COVID-19 Vaccine-Induced Myocarditis. Li X, editor. Adv Med [Internet]. 2024 Apr 18 [cited 2024 May 29];2024:1–17. Available from: https://www.hindawi.com/journals/amed/2024/4470326/
130. Wang YWY, Liu R Ben, Huang CY, Li HY, Zhang ZX, Li XZ, et al. Global, regional, and national burdens of myocarditis, 1990–2019: systematic analysis from GBD 2019: GBD for myocarditis. BMC Public Health [Internet]. 2023 Dec 1 [cited 2024 May 29];23(1):1–14. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-023-15539-5
131. Husby A, Hansen JV, Fosbøl E, Thiesson EM, Madsen M, Thomsen RW, et al. SARS-CoV-2 vaccination and myocarditis or myopericarditis: population based cohort study. BMJ [Internet]. 2021 Dec 1 [cited 2024 May 29];375. Available from: https://pubmed.ncbi.nlm.nih.gov/34916207/
132. Karlstad Ø, Hovi P, Husby A, Härkänen T, Selmer RM, Pihlström N, et al. SARS-CoV-2 Vaccination and Myocarditis in a Nordic Cohort Study of 23 Million Residents. JAMA Cardiol [Internet]. 2022 Jun 1 [cited 2024 May 29];7(6):600–12. Available from: https://pubmed.ncbi.nlm.nih.gov/35442390/
133. ACCESS-Background rate of adverse events-definition –Narcolepsy. [cited 2024 May 22]; Available from: https://zenodo.org/records/5226406
134. Silber MH, Krahn LE, Olson EJ, Pankratz VS. The Epidemiology of Narcolepsy in Olmsted County, Minnesota: A Population-Based Study. Sleep [Internet]. 2002 Jan 1 [cited 2024 May 27];25(2):197–202. Available from: https://dx.doi.org/10.1093/sleep/25.2.197
135. Cullen LA, Grange Z, Antal K, Waugh L, Alsina MS, Gibbons CL, et al. COVID-19 vaccine safety in Scotland – background rates of adverse events of special interest. Public Health. 2023 Nov 1;224:1–7.
136. Dodd CN, De Ridder M, Huang WT, Weibel D, Giner-Soriano M, Perez-Vilar S, et al. Incidence rates of narcolepsy diagnoses in Taiwan, Canada, and Europe: The use of statistical simulation to evaluate methods for the rapid assessment of potential safety issues on a population level in the SOMNIA study. PLoS One [Internet]. 2018 Oct 1 [cited 2024 May 27];13(10):e0204799. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0204799
137. Context | Pancreatitis | Guidance | NICE [Internet]. [cited 2024 May 10]. Available from: https://www.nice.org.uk/guidance/ng104/chapter/Context#acute-pancreatitis-2
138. Iannuzzi JP, King JA, Leong JH, Quan J, Windsor JW, Tanyingoh D, et al. Global Incidence of Acute Pancreatitis Is Increasing Over Time: A Systematic Review and Meta-Analysis. Gastroenterology [Internet]. 2022 Jan 1 [cited 2024 May 29];162(1):122–34. Available from: http://www.gastrojournal.org/article/S0016508521035460/fulltext
139. Li C li, Jiang M, Pan C qiu, Li J, Xu L gang. The global, regional, and national burden of acute pancreatitis in 204 countries and territories, 1990–2019. BMC Gastroenterol [Internet]. 2021 Dec 1 [cited 2024 May 29];21(1):1–12. Available from: https://bmcgastroenterol.biomedcentral.com/articles/10.1186/s12876-021-01906-2
140. AESI Case Definition Companion Guide: Myocarditis and Pericarditis [Internet]. [cited 2024 May 10]. Available from: https://zenodo.org/records/6668895
141. Andreis A, Imazio M, Casula M, Avondo S, Brucato A. Recurrent pericarditis: an update on diagnosis and management. Intern Emerg Med [Internet]. 2021 Apr 1 [cited 2024 May 29];16(3):551. Available from: /pmc/articles/PMC7914388/
142. Pericarditis epidemiology and demographics - wikidoc [Internet]. [cited 2024 May 29]. Available from: https://www.wikidoc.org/index.php/Pericarditis_epidemiology_and_demographics
143. Hadjadj J, Canzian A, Karadag O, Contis A, Maurier F, Sanges S, et al. Use of biologics to treat relapsing and/or refractory polyarteritis nodosa: data from a European collaborative study. Rheumatology [Internet]. 2022 Dec 23 [cited 2024 May 22];62(1):341–6. Available from: https://dx.doi.org/10.1093/rheumatology/keac332
144. De Virgilio A, Greco A, Magliulo G, Gallo A, Ruoppolo G, Conte M, et al. Polyarteritis nodosa: A contemporary overview. Autoimmun Rev. 2016 Jun 1;15(6):564–70.
145. Raval P. Polyarteritis. xPharm: The Comprehensive Pharmacology Reference. 2007 Jan 1;1–4.
146. ALAMANOS Y, VOULGARI P V., DROSOS AA. Incidence and Prevalence of Psoriatic Arthritis: A Systematic Review. J Rheumatol. 2008;35(7).
147. Queiro R, Loredo M, Braña I, Pardo E, Alonso S, Alperi M. Managing psoriatic arthritis in different clinical scenarios. Expert Rev Clin Immunol [Internet]. 2023 Dec 2 [cited 2024 May

14];19(12):1469–84. Available from: https://www.tandfonline.com/doi/abs/10.1080/1744666X.2023.2249235
148. Ravelli A, Martini A. Juvenile idiopathic arthritis. The Lancet [Internet]. 2007 Mar 3 [cited 2024 May 15];369(9563):767–78. Available from: http://www.thelancet.com/article/S0140673607603638/fulltext
149. Karmacharya P, Crowson CS, Bekele D, Achenbach SJ, Davis JM, Ogdie A, et al. The Epidemiology of Psoriatic Arthritis Over Five Decades: A Population-Based Study. Arthritis & Rheumatology [Internet]. 2021 Oct 1 [cited 2024 May 15];73(10):1878–85. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.41741
150. Sewerin P, Brinks R, Schneider M, Haase I, Vordenbäumen S. Prevalence and incidence of psoriasis and psoriatic arthritis. Ann Rheum Dis [Internet]. 2019 Feb 1 [cited 2024 May 15];78(2):286–7. Available from: https://ard.bmj.com/content/78/2/286
151. Hackett S, Coates L. Psoriatic arthritis: An up to date overview. Indian J Rheumatol [Internet]. 2020 Jun 1 [cited 2024 May 29];15(5):S45–51. Available from: https://journals.lww.com/ijru/fulltext/2020/15001/psoriatic_arthritis__an_up_to_date_overview.9.aspx
152. Gollamudi J, Sartain SE, Navaei AH, Aneja S, Kaur Dhawan P, Tran D, et al. Thrombosis and thromboembolism: Brighton collaboration case definition and guidelines for data collection, analysis, and presentation of immunization safety data. Vaccine [Internet]. 2022 Oct 19 [cited 2024 May 10];40(44):6431–44. Available from: https://pubmed.ncbi.nlm.nih.gov/36150973/
153. Germini F, Zarabi S, Eventov M, Turcotte M, Li M, de Wit K. Pulmonary embolism prevalence among emergency department cohorts: A systematic review and meta-analysis by country of study. Journal of Thrombosis and Haemostasis [Internet]. 2021 Jan 1 [cited 2024 May 10];19(1):173–85. Available from: http://www.jthjournal.org/article/S1538783622004354/fulltext
154. Bělohlávek J, Dytrych V, Linhart A. Pulmonary embolism, part I: Epidemiology, risk factors and risk stratification, pathophysiology, clinical presentation, diagnosis and nonthrombotic pulmonary embolism. Exp Clin Cardiol [Internet]. 2013 [cited 2024 May 29];18(2):129. Available from: /pmc/articles/PMC3718593/
155. Members AF, Torbicki A, Perrier A, Konstantinides S, Agnelli G, Galiè N, et al. Guidelines on the diagnosis and management of acute pulmonary embolism: The Task Force for the Diagnosis and Management of Acute Pulmonary Embolism of the European Society of Cardiology (ESC). Eur Heart J [Internet]. 2008 Sep 1 [cited 2024 May 29];29(18):2276–315. Available from: https://dx.doi.org/10.1093/eurheartj/ehn310
156. Rhabdomyolysis | PM&R KnowledgeNow [Internet]. [cited 2024 May 29]. Available from: https://now.aapmr.org/rhabomyolysis/
157. Stahl K, Rastelli E, Schoser B. A systematic review on the definition of rhabdomyolysis. J Neurol [Internet]. 2020 Apr 1 [cited 2024 May 10];267(4):877–82. Available from: https://pubmed.ncbi.nlm.nih.gov/30617905/
158. Farrukh AM, Quintanilla Cardona EA, Zakraoui RS, Artola Hernández YE, Vargas Vizcaya RE, Artola Hernandez RB, et al. Recent Advances and Updates in Rhabdomyolysis: A Comprehensive Review. JOJ Urology & Nephrology [Internet]. 2024 Jan 22 [cited 2024 May 29];8(4):1–6. Available from: http://juniperpublishers.com/jojun/JOJUN.MS.ID.555744.php
159. Stanley M, Chippa V, Aeddula NR, Rodriguez BSQ, Adigun R. Rhabdomyolysis. StatPearls [Internet]. 2023 Apr 16 [cited 2024 May 29]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK448168/
160. Cabral BMI, Edding SN, Portocarrero JP, Lerma E V. Rhabdomyolysis. Disease-a-Month. 2020 Aug 1;66(8):101015.
161. Safitri N, Alaina MF, Pitaloka DAE, Abdulah R. A Narrative Review of Statin-Induced Rhabdomyolysis: Molecular Mechanism, Risk Factors, and Management. Drug Healthc Patient Saf [Internet]. 2021 [cited 2024 May 29];13:211. Available from: /pmc/articles/PMC8593596/
162. Alpers JP, Jones LK. Natural history of exertional rhabdomyolysis: a population-based analysis. Muscle Nerve [Internet]. 2010 [cited 2024 May 29];42(4):487–91. Available from: https://pubmed.ncbi.nlm.nih.gov/20730874/
163. Aalborg C, Rød-Larsen C, Leiro I, Aasebø W. Økning i antall pasienter med treningsindusert rabdomyolyse? Tidsskrift for Den norske legeforening [Internet]. 2016 Oct 11 [cited 2024 May 22];136(18):1532–6. Available from: https://tidsskriftet.no/2016/10/originalartikkel/okning-i-antall-pasienter-med-treningsindusert-rabdomyolyse
164. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against

Rheumatism collaborative initiative. Arthritis Rheum [Internet]. 2010 [cited 2024 May 10];62(9):2569–81. Available from: https://pubmed.ncbi.nlm.nih.gov/20872595/
165. Rheumatoid Arthritis | Treatment & Management | Point of Care [Internet]. [cited 2024 May 22]. Available from: https://www.statpearls.com/point-of-care/17856
166. Chauhan K, Jandu JS, Brent LH, Al-Dhahir MA. Rheumatoid Arthritis. Rosen and Barkin's 5-Minute Emergency Medicine Consult: Fifth Edition [Internet]. 2023 May 25 [cited 2024 May 22]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK441999/
167. Lundkvist J, Kastäng F, Kobelt G. The burden of rheumatoid arthritis and access to treatment: Health burden and costs. European Journal of Health Economics [Internet]. 2008 Jan 21 [cited 2024 May 10];8(SUPPL. 8):49–60. Available from: https://link.springer.com/article/10.1007/s10198-007-0088-8
168. Myasoedova E, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising? Results from Olmsted County, Minnesota, 1955-2007. Arthritis Rheum. 2010 Jun;62(6):1576–82.
169. Galloway J, Capron JP, De Leonardis F, Fakhouri W, Rose A, Kouris I, et al. The impact of disease severity and duration on cost, early retirement and ability to work in rheumatoid arthritis in Europe: an economic modelling study. Rheumatol Adv Pract [Internet]. 2020 Jul 1 [cited 2024 May 10];4(2). Available from: https://dx.doi.org/10.1093/rap/rkaa041
170. Crowson CS, Matteson EL, Myasoedova E, Michet CJ, Ernste FC, Warrington KJ, et al. The lifetime risk of adult-onset rheumatoid arthritis and other inflammatory autoimmune rheumatic diseases. Arthritis Rheum. 2011 Mar;63(3):633–9.
171. Eriksson JK, Neovius M, Ernestam S, Lindblad S, Simard JF, Askling J. Incidence of rheumatoid arthritis in Sweden: A Nationwide population-based assessment of incidence, its determinants, and treatment penetration. Arthritis Care Res (Hoboken). 2013 Jun 1;65(6):870–8.
172. Safiri S, Kolahi AA, Hoy D, Smith E, Bettampadi D, Mansournia MA, et al. Global, regional and national burden of rheumatoid arthritis 1990–2017: a systematic analysis of the Global Burden of Disease study 2017. Ann Rheum Dis [Internet]. 2019 Nov 1 [cited 2024 May 10];78(11):1463–71. Available from: https://ard.bmj.com/content/78/11/1463
173. Rossini M, Rossi E, Bernardi D, Viapiana O, Gatti D, Idolazzi L, et al. Prevalence and incidence of rheumatoid arthritis in Italy. Rheumatol Int [Internet]. 2014 Mar 9 [cited 2024 May 28];34(5):659–64. Available from: https://link.springer.com/article/10.1007/s00296-014-2974-6
174. Duong TA, Valeyrie-Allanore L, Wolkenstein P, Chosidow O. Severe cutaneous adverse reactions to drugs. Lancet [Internet]. 2017 Oct 28 [cited 2024 May 10];390(10106):1996–2011. Available from: https://pubmed.ncbi.nlm.nih.gov/28476287/
175. Schreiber BE, Agrup C, Haskard DO, Luxon LM. Sudden sensorineural hearing loss. The Lancet [Internet]. 2010 Apr 3 [cited 2024 May 10];375(9721):1203–11. Available from: http://www.thelancet.com/article/S0140673609620717/fulltext
176. AESI Case Definition Companion Guide: Sensorineural Hearing Loss [Internet]. [cited 2024 May 10]. Available from: https://zenodo.org/records/7705371
177. Zanoni G, Girolomoni G, Bonetto C, Trotta F, Häusermann P, Opri R, et al. Single organ cutaneous vasculitis: Case definition & guidelines for data collection, analysis, and presentation of immunization safety data. Vaccine [Internet]. 2016 Dec 12 [cited 2024 May 10];34(51):6561–71. Available from: https://pubmed.ncbi.nlm.nih.gov/28029543/
178. Garcia-Porrua C, Gonzalez-Gay MA. Comparative clinical and epidemiological study of hypersensitivity vasculitis versus Henoch-Schönlein purpura in adults. Semin Arthritis Rheum. 1999 Jun 1;28(6):404–12.
179. Carlson JA, Ng BT, Chen KR. Cutaneous vasculitis update: Diagnostic criteria, classification, epidemiology, etiology, pathogenesis, evaluation and prognosis. American Journal of Dermatopathology [Internet]. 2005 Dec [cited 2024 May 29];27(6):504–28. Available from: https://journals.lww.com/amjdermatopathology/fulltext/2005/12000/cutaneous_vasculitis_update__diagnostic_criteria,.7.aspx
180. Ding B, Pignot M, Garal-Pantaler E, Villinger B, Schefzyk S, Desta B, et al. The Impact of Systemic Lupus Erythematosus Flares on Clinical and Economic Outcomes: The CHAMOMILE Claims Database Study in Germany. Rheumatol Ther [Internet]. 2024 Apr 1 [cited 2024 May 10];11(2):285–99. Available from: https://link.springer.com/article/10.1007/s40744-023-00635-0
181. Lupus Erythematosus, Systemic - MeSH - NCBI [Internet]. [cited 2024 May 10]. Available from: https://www.ncbi.nlm.nih.gov/mesh/68008180
182. Cornet A, Andersen J, Myllys K, Edwards A, Arnaud L. Living with systemic lupus erythematosus in 2020: a European patient survey. Lupus Sci Med [Internet]. 2021 Apr 1 [cited 2024 May 10];8(1):e000469. Available from: https://lupus.bmj.com/content/8/1/e000469

183. Tian J, Zhang D, Yao X, Huang Y, Lu Q. Global epidemiology of systemic lupus erythematosus: a comprehensive systematic analysis and modelling study. Ann Rheum Dis [Internet]. 2023 Mar 1 [cited 2024 May 10];82(3):351–6. Available from: https://ard.bmj.com/content/82/3/351
184. Wise RP, Bonhoeffer J, Beeler J, Donato H, Downie P, Matthews D, et al. Thrombocytopenia: Case definition and guidelines for collection, analysis, and presentation of immunization safety data. Vaccine. 2007 Aug 1;25(31):5717–24.
185. Voss EA, Shoaibi A, Yin Hui Lai L, Blacketer C, Alshammari T, Makadia R, et al. 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. EClinicalMedicine [Internet]. 2023 Apr 1 [cited 2024 May 28];58. Available from: https://pubmed.ncbi.nlm.nih.gov/37034358/
186. Xu S, Hong V, Sy LS, Glenn SC, Ryan DS, Morrissette KL, et al. Changes in incidence rates of outcomes of interest in vaccine safety studies during the COVID-19 pandemic. Vaccine. 2022 May 20;40(23):3150–8.
187. ACCESS-Background rate of adverse events-definition –transverse myelitis. [cited 2024 May 10]; Available from: https://zenodo.org/records/5237332
188. Burn E, Li X, Kostka K, Stewart HM, Reich C, Seager S, et al. Background rates of five thrombosis with thrombocytopenia syndromes of special interest for COVID-19 vaccine safety surveillance: Incidence between 2017 and 2019 and patient profiles from 38.6 million people in six European countries. Pharmacoepidemiol Drug Saf [Internet]. 2022 May 1 [cited 2024 May 10];31(5):495–510. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/pds.5419
189. Epidemiology of thrombotic thrombocytopenia syndrome in integrated health-care database in England Secondary data analysis using a cohort design (Epidemiology of TTS) | HMA-EMA Catalogues of real-world data sources and studies [Internet]. [cited 2024 May 29]. Available from: https://catalogues.ema.europa.eu/node/3788/administrative-details
190. Feuerstein JD, Moss AC, Farraye FA. Ulcerative Colitis. Mayo Clin Proc [Internet]. 2019 Jul 1 [cited 2024 May 10];94(7):1357–73. Available from: http://www.mayoclinicproceedings.org/article/S0025619619300849/fulltext
191. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. The Lancet [Internet]. 2017 Dec 23 [cited 2024 May 10];390(10114):2769–78. Available from: http://www.thelancet.com/article/S0140673617324480/fulltext
192. Cosnes J, Gowerrousseau C, Seksik P, Cortot A. Epidemiology and Natural History of Inflammatory Bowel Diseases. Gastroenterology [Internet]. 2011 May 1 [cited 2024 May 10];140(6):1785-1794.e4. Available from: http://www.gastrojournal.org/article/S0016508511001648/fulltext
193. Trandafir LM, Spoiala EL, Ghiga G, Gimiga N, Budescu PD, Lupu VV, et al. Impact of COVID-19 on Pediatric Inflammatory Bowel Diseases—From Expectations to Reality. Journal of Personalized Medicine 2024, Vol 14, Page 399 [Internet]. 2024 Apr 9 [cited 2024 May 28];14(4):399. Available from: https://www.mdpi.com/2075-4426/14/4/399/htm
194. Malik A, Ha NB, Barnes GD. Choice and Duration of Anticoagulation for Venous Thromboembolism. Journal of Clinical Medicine 2024, Vol 13, Page 301 [Internet]. 2024 Jan 4 [cited 2024 May 28];13(1):301. Available from: https://www.mdpi.com/2077-0383/13/1/301/htm
195. Pastori D, Cormaci VM, Marucci S, Franchino G, Del Sole F, Capozza A, et al. A Comprehensive Review of Risk Factors for Venous Thromboembolism: From Epidemiology to Pathophysiology. International Journal of Molecular Sciences 2023, Vol 24, Page 3169 [Internet]. 2023 Feb 5 [cited 2024 May 28];24(4):3169. Available from: https://www.mdpi.com/1422-0067/24/4/3169/htm
196. Næss IA, Christiansen SC, Romundstad P, Cannegieter SC, Rosendaal FR, Hammerstrøm J. Incidence and mortality of venous thrombosis: a population-based study. Journal of Thrombosis and Haemostasis. 2007 Apr 1;5(4):692–9.
197. Xie J, Prats-Uribe A, Feng Q, Wang Y, Gill D, Paredes R, et al. Clinical and Genetic Risk Factors for Acute Incident Venous Thromboembolism in Ambulatory Patients With COVID-19. JAMA Intern Med [Internet]. 2022 Oct 1 [cited 2024 May 28];182(10):1063–70. Available from: https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2795466
198. Cushman M, Tsai AW, White RH, Heckbert SR, Rosamond WD, Enright P, et al. Deep vein thrombosis and pulmonary embolism in two cohorts: The longitudinal investigation of thromboembolism etiology. American Journal of Medicine [Internet]. 2004 Jul 1 [cited 2024 May 28];117(1):19–25. Available from: http://www.amjmed.com/article/S0002934304001676/fulltext

199. Lutsey PL, Zakai NA. Epidemiology and prevention of venous thromboembolism. Nature Reviews Cardiology 2022 20:4 [Internet]. 2022 Oct 18 [cited 2024 May 28];20(4):248–62. Available from: https://www.nature.com/articles/s41569-022-00787-6