SAFETY-VAC: A European Framework for Post-Authorisation Vaccine Safety MonitoringUsing Real-World Data

Sima Mohammadi[1], Carlos E. Durán[1], Rosa Gini[2], Davide Messina[2], Fabio Riefolo[3], Judit Riera, Vjola Hoxha[1],Irene Pazos[3], Vera Ehrenstein[4], Giuseppe Roberto[2], Miriam Sturkenboom[1]

1. Department of Data Science & Biostatistics, Julius Global Health, University Medical Center Utrecht (UMCU), Utrecht, the Netherlands
2. Agenzia Regionale di Sanitá Toscana (ARS) Data Tuscany region
3. Teamit Institut
4. Department of Clinical Epidemiology, Aarhus University and Aarhus University Hospital

UMC Utrecht

Objective

To describe and assess the fitness-for-purpose of a European network of real-world data (RWD) sources for conducting vaccine safety studies.

Introduction

- While vaccine safety is rigorously evaluated in pre-authorisation clinical trials, new or rare safety concerns may arise only after widespread use. The SAFETY-VAC project supports the Vaccine Monitoring Platform established by the European Medicines Agency (EMA) and the European Centre for Disease Prevention and Control (ECDC) to ensure timely, data-driven post-authorisation safety evaluation of vaccines in the European Union (EU) and European Economic Area (EEA).

Methods

We conducted a retrospective, population-based cohort study using electronic health records and registries from 10 data sources across 7 European countries (Spain, Italy, Denmark, Norway, Finland, France, and the United Kingdom), covering over 53 million individuals from January 2017 to the most recent data availability. Data were transformed into the ConcePTION Common Data Model (CDM), and quality was evaluated using the INSIGHT R tool (Levels 1–3 (Figure.1)). Study variables included vaccine exposures, 39 pre-specified adverse events of special interest (AESIs), and key covariates. Vaccine coverage, event incidence rates, and covariate prevalence were benchmarked against World Health Organization (WHO)/ECDC indicators. Fitness-for-purpose was assessed using the Structured Process to Identify Fit-For-Purpose Data (SPIFD) framework.

Results

Nine data sources passed Level 1 and Level 2 data quality checks; seven also completed the analytical scripts. All eligible databases captured vaccine exposures via ATC classification or mapped vaccine types. Coverage estimates for childhood, adolescent, and COVID-19 vaccines were derived using inverse probability weighting and period prevalence follow-up methods. For most childhood vaccines, BIFAP, SIDIAP, and NHR estimates aligned closely with WHO targets. COVID-19 vaccine coverage stratified by age and dose was consistent with ECDC data, particularly in individuals aged ≥60.

Incidence rates for AESIs such as myocarditis, Bell's palsy, cerebral venous sinus thrombosis, and autoimmune hepatitis were largely consistent with published estimates. Discrepancies observed for sensorineural hearing loss, rheumatoid arthritis, and pancreatitis suggest possible misclassification or setting-specific variation. Most rates were comparable to values from the COVID-19 Vaccine Monitoring (CVM) and ACCESS studies. Covariate prevalence, including cardiovascular disease and diabetes, reflected expected population distributions. Data completeness varied across vaccines and outcomes (Table 1) . Three data sources could not complete analysis due to delays in extraction or approvals.

The SPIFD evaluation identified seven data sources (BIFAP, SIDIAP, VID, PEDIANET, CPRD, DHR, and NHR) as fit-for-purpose for rapid vaccine safety surveillance (Table 2).

Figure 1. Quality checks' INSIGHT tool steps and criteria

INSIGHT Quality control framework		
Level 1 (Completeness of data)	Level 2 (logical relationship and integrity of values within and between variables and tables)	Level 3 (Data content characterization)
1. Birth date elements	1. Date values before birth	1. Population tree as expected for the data source
2.Person with a recorded death date	2. Date values after death	2. Temporal patterns in Event rates
3. Sex recorded	3. Date values outside observation periods	3. Level of detail of ATC codes
4. Recorded country of birth	4. Person IDs not in PERSONS table	4. Temporal patterns in medicines rates
5. Recorded events codes	5. CDM table with highest percentage of issues	5. Vaccination exposure identification
6. Recorded codes for medicines records		6. Temporal patternsin vaccination rates
7. Recorded codes to identify vaccine		
8. Vaccine dose recorded		

Table 1. Data sources characteristics

Data Access provider (Data source)	Data banks available for SAFTEY_VAC study
AEMPS (BIFAP)	Primary care, hospital discharge diagnosis, community pharmacy dispensing, date of death.
IDIAP J Gol (SIDIAP)	Primary care, outpatient specialist, outpatient laboratory results, surveillance data, EMR visits, hospital discharge diagnosis, long term facility diagnosis, date of death.
FISABIO (VID)	Primary care , outpatient specialist , outpatient laboratory results, surveillance data, EMR visits, hospital discharge diagnosis, inhospital/outpatinet prescribing, long term facility diagnosis, date and reasons of death.
IACS (EPICHRON)	Primary care, outpatient laboratory results, EMR visits, hospital discharge diagnosis, pharmacy dispensing outpatient, date of death.
SOSETE (PEDIANET)	Primary care , outpatient specialist diagnosis, surveillance data, EMR visits, hospital discharge diagnosis, inhospital/outpatinet prescribing , date and reason of death.
Utrecht University (CPRD-Aurum)	Primary care diagnoses, prescriptions, lab tests, hospital admissions and procedures CPRD death date
Aarhus University (Danish registries)	Outpatient specialist diagnosis, inhospital laboratory results, EMR visits, hospital discharge diagnosis, inhospital/outpatient prescribing, date of death, reasons of death (2 years lag time)
University of Eastern Finland (Finnish registries)	Primary care record (with some restrictions), outpatient specialist diagnosis,surveillance data, EMR visits, hospital discharge diagnosis,outpatient and inhospital laboratory results, outpatient pharmacy dispensing, long term facility diagnoses, date and reasons of death.
BPE (SNDS)	Outpatient healthcare (no results, no indication), pharmacy dispensing (quantity, dosage, name), public/private hospital stays with discharge diagnosis (no results), public hospital visits (no results, no indication), EMR visits (with diagnosis if > 1 day, without if <=1 day), in-hospital dispensing/prescription (only for outof-DRG drugs), date and reason of death
University of Oslo (Norwegian registries)	Primary care record, outpatient specialist diagnosis, surveillance data (infectious diseases), EMR visits, hospital discharge diagnosis, outpatient pharmacy dispensing, date and reasons of death.

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
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 diagnosis codes and drug proxies)	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
	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
	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

Conclusions

SAFETY-VAC demonstrates the feasibility of leveraging a federated RWD network to support post-authorisation vaccine safety surveillance. Data quality and completeness across key variables were sufficient in most sources, though updates to data instances and full ETL of childhood vaccine data remain necessary for optimal readiness. The SPIFD tool provided a transparent framework for assessing data readiness. Continued updates and standardization of data sources are essential for sustaining rapid-response capacity in future safety evaluations.