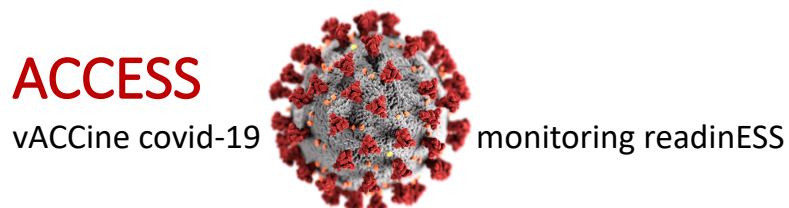


Coverage of COVID-19 vaccines in electronic healthcare databases: a protocol template from the ACCESS project



This protocol can be used by organizations to monitor COVID-19 vaccines post-introduction.
Please reference as

Sturkenboom M. et al. Coverage evaluation of COVID-19 vaccines in electronic healthcare databases: a protocol template from the ACCESS project

DISCLAIMER

This protocol has been accepted by EMA as a deliverable of the framework contract No EMA/2018/28/PE. The protocol expresses the expertise of the authors and the ACCESS consortium as well as feedback received from EMA. It may not be understood or quoted as being made on behalf, or reflecting the position of the European Medicines Agency or one of its Committees or Working Parties

Study Information

Title	Coverage evaluation of COVID-19 vaccines in immunization registries and/or health care databases: a protocol template from the ACCESS project
Protocol version identifier	0.2
Date of last version of protocol	28 January 2021
EU PAS Register number	<i>Registration number in the EU PAS Register; indicate "Study not registered" if the study has not been registered in the EU PAS Register.</i>
Active substance	<i>COVID-19 vaccines (J07)</i>
Medicinal product	<i>List of centrally authorised medicinal product(s) and/or, if possible, of nationally authorised products subject to the study</i>
Product reference	<i>Reference number(s) of centrally authorised products and/or, if possible, of nationally authorised products subject to the study</i>
Procedure number	<i>If applicable, Agency or national procedure number(s), e.g., EMA/X/X/XXX</i>
Marketing authorization holder(s)	<i>Marketing authorisation holder(s) which initiate(s), manage(s) or finance(s) the study</i>
Research question and objectives	To evaluate the coverage of specific COVID-19 vaccines
Country(-ies) of study	Eligible data access providers
Authors	Miriam Sturkenboom, University Medical Centre Utrecht
Other keys contributors	

Marketing authorisation holder(s)

Marketing authorisation holder(s)	<i>Name, address and contact details of the marketing authorisation holder(s).</i>
MAH contact person	<i>Contact person for this PASS protocol submission (if this a joint PASS, only one person should be mentioned)</i>

Trademarks

Brand Name	Generic Name	Trademark Holder

1 Table of Contents

1 Table of Contents.....	4
2 List of Abbreviations.....	6
3 Responsible Parties.....	7
4 Abstract.....	8
5 Amendments and Updates	9
6 Milestones and Timeline	9
7 Rationale and Background	10
8 Research Question and Objectives	11
9 Research Methods.....	11
9.1 Study Design	12
9.2 Setting	12
9.2.1 Source Population	12
9.2.2 Study Period, Population and Follow-up Period.....	12
9.2.3 Variables.....	12
9.3 Data Sources	16
9.4 Study Size	16
9.5 Data Collection and Management.....	16
9.6 Data Analysis.....	18
9.6.1 Descriptive Analysis.....	18
9.6.2 Measures of coverage.....	19
Period Prevalence (PP)	20
Period Prevalence: Follow-Up (PPFU).....	20
Cumulative distribution function (CDF).....	20
9.6.3 Data Integration	22
9.6.4 Subgroup Analysis	22
9.6.5 Sensitivity Analysis	22
9.7 Quality Control	23
9.8 Limitations of the Research Methods	23

9.9 Other Aspects	24
10 Protection of Human Subjects	24
11 Management and Reporting of Adverse Events /Adverse Reactions	24
12 Plans for Disseminating and Communicating Study Results	25
13 Other Good Research Practice.....	25
14 References	27
Annex 1. List of Stand-Alone Documents	29
Annex 2. ENCePP Checklist for Study Protocols.....	30

2 List of Abbreviations

ACCESS	vACCine covid-19 monitoring readinESS
ADVANCE	Accelerated Development of VAccine beNefit-risk Collaboration in Europe
CDC	Centres for Disease Control and Prevention
CDM	Common Data Model
DAP	Data Access Provider
ECDC	European Centre for Disease Prevention and Control
EMA	European Medicines Agency
EMR	Electronic Medical Records
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance.
EU PAS	The European Union electronic Register of Post-Authorisation Studies
GDPR	General Data Protection Regulation
GPP	Good Participatory Practice
ICD	International Classification of Diseases
ICMJE	International Committee of Medical Journal Editors
IMI	Innovative Medicines Initiative
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2

3 Responsible Parties

To be completed by study investigator(s) when they use the protocol

[Principal investigator institution name]
Address

name, degrees, job title

name, degrees, job title

name, degrees, job title

name, degrees, job title

name, degrees, job title

name, degrees, job title

[Sponsor name]
Address

name, degrees, job title

name, degrees, job title

name, degrees, job title

name, degrees, job title

name, degrees, job title

name, degrees, job title

Collaborating Institutions

Study Sites

4 Abstract

This section should be filled out with the following information.

Title:

Rationale and background:

Research question and objectives:

Study design:

Population:

Variables:

Data sources:

Study size:

Data analysis:

Milestones:

5 Amendments and Updates

<<to be filled upon actual protocol>>

6 Milestones and Timeline

This section should be filled out with the following information, when the study is implemented.

Milestone	Date
Start of data collection	
End of data collection	
<Study progress report(s) 1>	
<Study progress report(s) 2>	
<Study progress report(s) 3>	
<Interim report 1>	
<Interim report 2>	
<Interim report 3>	
<Registration in the EU PAS Register>	
Final report of study results	

EMA = European Medicines Agency; EU PAS Register = European Union Electronic Register of Post-Authorisation Studies.

7 Rationale and Background

The novel coronavirus SARS-CoV-2, the cause of COVID-19, has led to a global pandemic. Several COVID-19 vaccine candidates are currently under research and in development. COVID-19 vaccines may be licensed by the European Medicines Agency (EMA) following what is likely to be an accelerated investigational and licensing procedure. Because the pre-licensure period is short and number of participants in clinical studies is limited, monitoring of the safety of vaccines in the post-introduction phase will be needed in an efficient manner.

Based on recent European Commission communication to the European Parliament¹ there were three contracts² in place in October 2020 that allowed the European Commission (EC) purchase of a COVID-19 vaccine once it has proven safe and effective, namely with AstraZeneca, Sanofi-GSK and Johnson & Johnson. As of November 2020, the Commission continued discussing similar agreements with other vaccine manufacturers (CureVac, Moderna & Pfizer). On Nov 11th, 2020, the commission also secured a contract with Pfizer, on Nov 17th with Curevac and on Nov 25th with Moderna³

The Commission has thus far secured access to the following doses of COVID-19 vaccines for Europe:

- AstraZeneca: 300 million doses.
- Sanofi-GSK: a purchase option for 300 million doses.
- Johnson & Johnson: 200 million doses.
- Pfizer 200 million doses & optional 100 million doses
- Moderna: 80 million doses and optional extra 80 million.
- Curevac: 225 million doses plus an option to request up to a further 180 million doses

As of January 2021, Pfizer and Moderna vaccines have been licensed. It is unknown whether all vaccines, will successfully complete the development and authorisation process and thus meet efficacy and safety criteria to be placed on the EU market.

An allocation methodology agreed between the Commission and Member States⁴, ensures that all Member States will have equal access to the available doses based on their population size. Once available and authorised at EU level, all Member States will have access to COVID- 19 vaccines at the same time. The overall number of vaccine doses will be limited during the initial stages of deployment and before production can be ramped up.

Meanwhile, high on the list of actions is a decision which groups should have priority access to vaccines. The WHO-SAGE committee developed a values framework for the allocation and prioritization of COVID-19

¹ Communication from the Commission to the European Parliament and the Council. Preparedness for COVID-19 vaccination strategies and vaccine deployment. 15 Oct 2020. Available at: https://ec.europa.eu/health/sites/health/files/vaccination/docs/2020_strategies_deployment_en.pdf

² On 14 August, the Commission reached a first agreement with the pharmaceutical company AstraZeneca to purchase 300 million doses of a potential vaccine against COVID-19. On 18 September, a second contract with Sanofi-GSK was signed, for an option that will allow all Member States to purchase up to 300 million doses of the Sanofi-GSK vaccine. On 8 October, the Commission approved an advance purchase agreement with Pharmaceutica NV, one of the Janssen Pharmaceutical Companies of Johnson & Johnson, allowing Member States to purchase vaccines for 200 million people.

³ Agreed by the Commission and Member States in the Agreement on the joint EU approach to COVID-19 vaccines procurement adopted by the Commission on 17 June and endorsed by all Member States.

⁴ WHO SAGE values framework for the allocation and prioritization of COVID-19 vaccination. 13 Sep 2020. Available at: https://www.nitag-resource.org/sites/default/files/2020-09/WHO-2019-nCoV-SAGE_Framework-Allocation_and_prioritization-2020.1-eng.pdf

vaccination⁴. The Framework articulates the overall goal of COVID-19 vaccine deployment, provides six core principles that should guide distribution and twelve objectives that further specify the six principles (human well-being, equal respect, global equity, national equity, reciprocity, legitimacy). To provide recommendations for allocating vaccines between countries and prioritizing groups for vaccination within each country, the Values Framework needs to be complemented with information about specific characteristics of available vaccine or vaccines, the benefit-risk assessment for different population groups, the amount and pace of vaccine supply, and the current state of the epidemiology, clinical management, and economic and social impact of the pandemic. Hence, the final vaccination strategy will be defined by the characteristics of vaccine products as they become available (2). In Europe, ECDC is convening the National Immunization Technical Advisory Group (NITAG) collaboration to guide and harmonize decision making in roll out of vaccines across EU countries. Based on the SAGE values framework and the evidence based on the ongoing clinical trials, the following groups seem likely to be addressed first, but this will be decided in each country separately in the coming months:

- Health care workers.
- Populations with significantly elevated risk of severe disease or death:
 - Older adults defined by age-based risk - may vary by country/region.
 - Older adults in high risk living situations (examples: long term care facility, those unable to physically distance).
 - Groups with comorbidities or health states (e.g. pregnancy/lactation) determined to be at significantly higher risk of severe disease or death.
 - Sociodemographic groups at disproportionately higher risk of severe disease or death.

As part of the preparedness activities for surveillance of COVID-19 vaccines, this template protocol provides a template for quickly developing a full study protocol to perform vaccine coverage studies through the secondary use of electronic healthcare databases and/or immunization registers.

The ACCESS project has developed several template protocols, that address: vaccine coverage, vaccine effectiveness and vaccine safety.

8 Research Question and Objectives

Primary objective:

To determine exposure and coverage to <<COVID-19 vaccine>>

Secondary objective:

To determine exposure and coverage to <<COVID-19 vaccine>> in specific subgroups that are targeted for vaccination

9 Research Methods

9.1 Study Design

Retrospective (multi)-database cohort study.

9.2 Setting

Population based data sources such as electronic medical record databases, claims databases or population-based immunizations registers that capture electronic information on the vaccine administered and the population that is underlying this.

9.2.1 Source Population

The source population for each of the study design will comprise all individuals registered in the data source during the study period for that data instance.

9.2.2 Study Period, Population and Follow-up Period

The study population will comprise all persons in the source population that are eligible for the study according to specific inclusion and exclusion criteria.

Study period should start upon COVID-19 vaccine introduction

Required Inclusion criteria comprise:

- Persons registered in the data source at a certain date
- Available information on date of birth and sex (If needed certain age categories can be specified)
- Available information on current eligibility (meaning knowing that person is in follow-up and did not leave/die)

Eligible individuals should be identified in each of the database using a pre-specified selection process and/or by applying pre-specified algorithms.

Follow-up time should start at the moment that the latest of the inclusion criteria is met.

follow-up ends at the earliest of the occurrence of censoring conditions or the last data draw down/data availability.

9.2.3 Variables

9.2.3.1 Exposure Assessment

In this section, operational definitions for identifying exposures to COVID-19 vaccines should be described, including the manners to identify them in the specific data sources.

Exposure of key interest is the specific COVID-19 vaccine brand and if possible lot number.

Definitions: Exposure is the number of doses of specific vaccines that are provided to the population, coverage is the percentage of the target population that has been vaccinated at a given moment.

Currently it is still not exactly known in each of the countries how vaccines will be rolled out, who will vaccinate and how the immunization will be captured. UK provides immunization cards to citizens that captures the brand a lot number.

Each secondary package of COVID-19 vaccine will be labelled with a DataMatrix⁵, which contains a GTIN code that is unique for pharmaceutical products.

“In supply chain management for medicinal products, the dominant standard is using one 2D symbology (GS1 DataMatrix), which should be applied on the secondary packaging (carton boxes), but can also be applied on the primary packaging (vial or prefilled syringe)” (Fig. 1).

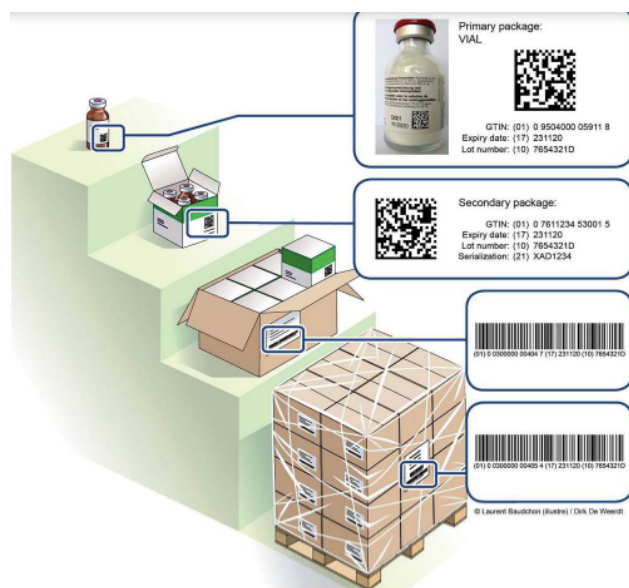


Figure 1: Representation of Global Trade Item Number, expiry date, lot number and serialisation number on different levels of vaccine packaging with two-dimensional DataMatrix and linear barcode (from van der Stichele et al. Vaccine 2020)

⁵ Vander Stichele RH, Hay C, Fladvad M, Sturkenboom MCJM, Chen RT. How to ensure we can track and trace global use of COVID-19 vaccines? [published online ahead of print, 2020 Nov 24]. Vaccine. 2020;S0264-410X(20)31508-5. doi:10.1016/j.vaccine.2020.11.055

Legend: The Global Trade Item Numbers differ on each level of packaging. Expiry date and lot number are the same on each level of packaging. Only the secondary package carries a fourth unique check (the serialisation code as a protection against falsification). Numbers are represented on primary and secondary package by GS1 Data Matrix for dispensing monitoring and by linear barcodes on higher levels of packaging for production and distribution monitoring.

We recommend that for assessing exposure in immunization registers the GTIN is used and linked to national portfolios that indicate the specific medicinal product, nationally

Countries will have to explore what information is recorded in the immunization register, and moreover if the Data Matrix can be scanned at immunization, which is challenging as the primary package does not need to be barcoded and the secondary package is bulky. Dedicated instructions need to be provided that immunizers can record the vaccine and link it to the individual.

Vaccine information should be obtained from immunization registers or other electronic sources that capture COVID-19 vaccine such as pharmacy dispensing records, general practice records, immunisation registers, medical records, or other secondary data sources.

If the vaccine is administered in multiple doses, doses should be classified separately. If the vaccine is administered as a single dose, then only the first exposure code during the study period will be considered, which assumes that the remaining codes could be recording errors and/or medication errors. In that case, a censoring at subsequent doses will be applied.

Operationalization

Vaccinations need to be obtained from the databases by using names of vaccines and database specific codes. Vaccines will be categorized into vaccine types by platform. Brand data will be obtained from the recorded data where available. Variables in a common vaccine input file are: coded patient identifier to link with population, date of administration, vaccine type (platform), brand, recorded dose (if available), derived dose (imputed based on chronological order and age/timing of administration) and lot number.

9.2.3.2 Covariate Assessment

Populations may be targeted for vaccination based on at risk conditions, age, profession or living situations. Based on the SAGE recommendations, each country makes its own decisions. At risk medical conditions for developing severe COVID-19 have been defined based on scientific evidence available on Centre Disease Control (CDC website, July 2020) and National Health Services (NHS website, July 2020) websites. Those websites are updated regularly and provide a classification of at-risk conditions for developing severe COVID-19 based on level of evidence.

The selected at-risk medical conditions are considered as at higher risk to develop severe COVID-19 (Table 1).

The following variables should be created:

- At-Risk groups: medical codes and associated dates for at-risk medical conditions characterizing at-risk groups for developing severe COVID-19 as well as prescription and/or dispensing records for drug exposures which may be used as proxies for their identification. At-risk groups will be created for each of the at-risk medical conditions listed in Table 1. Multimorbidity will not be considered (subjects may belong to more than one at-risk group).

- Pregnancy start and end dates: pregnancy start and end dates will be assessed from medical birth registers for those data sources with access to a registry, while existing algorithms for defining start and end of pregnancy will be utilized in those data sources with an existing algorithm.

Table 1: At risk medical conditions and drug proxies from electronic health care data

At-risk medical conditions	Medicinal product proxy(ies) (ATC code)
Cancer (with chemo/immuno/radio-therapy, cancer treatment, immunosuppressant; targeted cancer treatment (such as protein kinase inhibitors or PARP inhibitors); blood or bone marrow cancer (such as leukaemia, lymphoma, myeloma))	Alkylating agents (L01A) Antimetabolites (L01B) Plant alkaloids and other natural products (L01C) Cytotoxic antibiotics and related substances (L01D) Other antineoplastic agents (L01X) Hormones and related agents (L02A) Hormone antagonists and related agents (L02B) Immunostimulants (L03) Immunosuppressants (L04)
Type 2 Diabetes	Blood glucose lowering drugs, excluding insulins(A10B)
Obesity (BMI > 30)	Peripherally acting antiobesity products (A08AB) Centrally acting antiobesity products (A08AA)
Cardiovascular disease/ Serious heart conditions including heart failure, coronary artery disease, cardiomyopathies	Antiarrhythmics, class I and III (C01B) Cardiac stimulants excl. Cardiac glycosides (C01C) Vasodilators used in cardiac diseases (C01D) Other cardiac preparations (C01E) Antithrombotic agents (B01A)
Chronic lung disease including COPD, cystic fibrosis, severe asthma	Drugs for obstructive airway diseases (R03) Lung surfactants (R07AA) Respiratory stimulants (R07AB)
Chronic kidney disease	Erythropoietin (B03XA01)
HIV	Protease inhibitors (J05AE) Combinations to treat HIV (J05AR) NRTI (J05AF) NNRTI (J05AG)
Immunosuppression	Immunosuppressants (L04A) Corticosteroids (H02)
Sickle Cell Disease	Hydroxyurea (L01XX05) Other haematological agents (B06AX)
Pregnancy	
Start date of pregnancy	

End date of pregnancy

Medical diagnosis codes for the covariates can be obtained from the ACCESS background rate protocol codesheets.

NOTE for writers of the protocol: Other risk groups should be defined nationally based on vaccination policy.

9.2.3.4.3 Effect Modifiers

In this section, any factors (e.g., age, chronic conditions, profession, living conditions) that are hypothesised to alter coverage of <<COVID-19 vaccine>> on the event of interest should be listed.

[Investigator should specify the effect modifiers and their operational definitions in this section].

9.3 Data Sources

The data sources for the exposures, and covariates should be listed in this section, including coding systems, data lag, and starting date of data availability. The size of the database and COVID-19 vaccination strategy should be included in this section. Validation data on relevant data elements should also be described.

9.4 Study Size

[To be completed by the study investigator(s) based on assumptions of number of cases and effect estimate size at the time of the full protocol development]

9.5 Data Collection and Management

This section assumes the approach of a distributed network of DAPs who agree to use a common protocol, common data model and common analytics. We recommend to prepare for a model where original data remain local, and are transformed in a common data model that will allow for study specific structural and semantic harmonization (model C, **Figure 2**) (Gini et al, 2020).

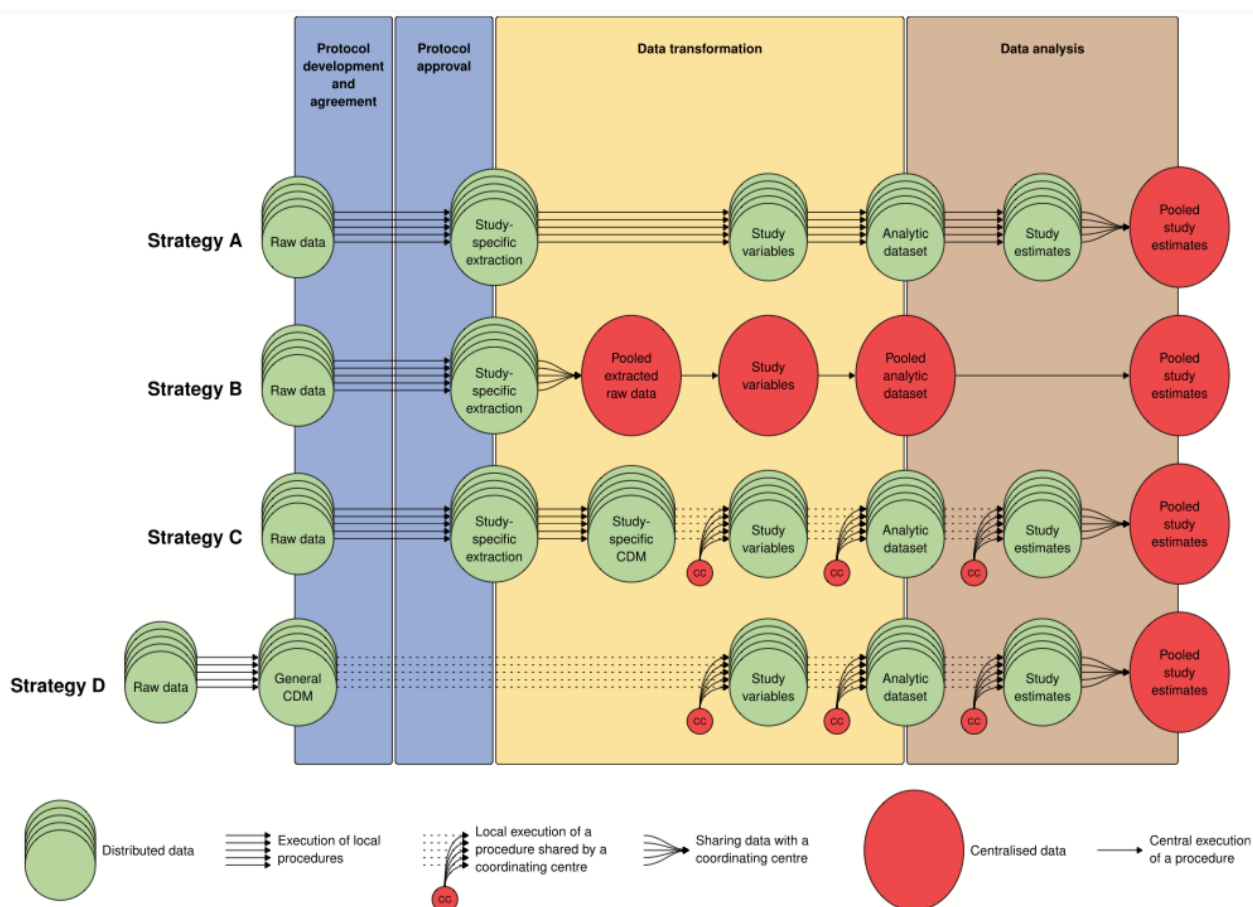


Figure 2. Options for multi-database studies in Europe

[If other models are used this section should be adapted]

In short, model C requires that each data access provider will extract the data required for the study and transform their local patient level data into a common data model (CDM). An example of a widely used common data model in Europe (currently 24 DAPs) is the ConcePTION CDM, which is publicly available⁶. Extract, transform and load (ETL) design and instructions are available, as well as tools to check the quality of the data for the AESI listed above as these are also utilized for the ACCESS background rate protocol and the coverage studies in the ADVANCE proof of concept studies.

A common program to run quality checks, data transformation, and analysis should be prepared and verified and be sent to all DAPs. Aggregate results and summary estimates resulting from the programs should be returned to a single coordinating centre for pooled meta-analysis and reporting.

⁶ <https://www.imi-conception.eu/wp-content/uploads/2020/10/ConcePTION-D7.5-Report-on-existing-common-data-models-and-proposals-for-ConcePTION.pdf>

Example of cleaning steps of vaccination data in the ADVANCE system testing are^{7 8}

- Removing observations with missing date at vaccination
- MinimumNecessaryDistance (distance between subsequent doses of the same vaccine) is implemented and allows for deleting doses (P1 first dose, P2 second dose) of the same vaccine that were administered too shortly after each other. (That is, if $\text{Date}_{P2} < \text{Date}_{P1} + \text{MinimumNecessaryDistance}$, then P2 is deleted.). The value is set to 1 day.
- In case of multiple records with the same value for the recorded dose and derived dose (depending on database), the one with the earliest date is kept, the other records are deleted. A new variable will be created (DoseCombined)
- In case of multiple records with the same date at vaccination but different values for dose, the record with earliest dose is kept; the other records are deleted. In case the dates of the ordered values for dose are not in a chronological order, the dates are swapped to the right chronological order.
- Merging the population and vaccination input files. Only records for patients in the population file are retained

[Investigators should modify this section as needed for the specific study; if specific procedures of the identified research partners are known, they can be included here]

Routine procedures should include checking electronic files, maintaining security and data confidentiality, following analysis plans, and performing quality-control checks of all programs. Each partner should maintain any patient-identifying information securely on site according to internal standard operating procedures or guidance documents.

Security processes should be in place to ensure the safety of all systems and data. Every effort should be made to ensure that data are kept secure so that they cannot be accessed by anyone except the study team.

Appropriate data storage and archiving procedures should be followed by each DAP and the coordinating organization, with periodic backups. Standard procedures should be in place at each research centre to restore files in the event of a hardware or software failure.

9.6 Data Analysis

9.6.1 Descriptive Analysis

Attrition diagrams demonstrating the loss of subjects applying inclusion and exclusion criteria should be provided.

⁷⁷ Sturkenboom M, van de Aa L, Bollaerts K, Emborg, HD; Gini R; Liyanage H; de Lusignan S; Stergioulas L; Khromava A, Switzer C, Ferreira G, Titievsky L, Weibel D. The ADVANCE system for evidence generation on vaccine coverage, benefits and risks based on secondary use of electronic health care data in Europe. Vaccine (submitted)

⁸ Emborg HD et al. ADVANCE system testing: can coverage of pertussis vaccination be estimated in European countries using electronic healthcare databases: an example. Vaccine (submitted)

Demographic characteristics of the study population (e.g. age at study entry, sex) and baseline characteristics (eg. co-morbidities) should be summarized for each data source using descriptive statistics. Description of <<COVID-19 vaccine>> type should be described and along with counts of exposure by calendar month to allow for inspection of time trends.

The distributions of the characteristics by <<COVID-19 vaccine>> should be presented in tables to demonstrate potential differences/or lack thereof between groups.

Counts and percentages should be presented for categorical variables (age at study entry in categories, sex, race/ethnicity, comorbidities). Mean, standard error, median and range should be presented for continuous variables (age at study entry). The missingness of variables should also be described.

Appendices should provide code /algorithm counts for the co-morbid conditions.

9.6.2 Measures of coverage

Based on the methodological developments in the ADVANCE project⁹, which aimed at assessing coverage from dynamic cohorts the following outcome parameters may be estimated:

- Number of persons and person-time by database
- Number of doses administered by vaccine and brand during study period

Coverage curves by database, by at risk groups, or by birthyears for vaccines in elderly, per dose. Estimates should be provided at certain ages based on cumulative distribution functions.

For the calculation of coverage, we recommend the following calculations, but selection is possible based on specific data characteristics⁹.

We assign a letter (A, B, C, D, E) to every age week 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 birthyear, we produce the following estimators;

⁹ Braeye T, Bauchau V, Sturkenboom M, Emborg HD, García AL, Huerta C, Merino EM, Bollaerts K. Estimation of vaccination coverage from electronic healthcare records; methods performance evaluation - A contribution of the ADVANCE-project. PLoS One. 2019 Sep 18;14(9):e0222296. doi: 10.1371/journal.pone.0222296. PMID: 31532806; PMCID: PMC6750592.

Period Prevalence (**PP**)

The period prevalence estimate for age i is the proportion of vaccinated persons over the total number of eligible persons. In other words; PP_i represents the cumulative incidence up to age i over all eligible persons in the cohort.

$$PP_i = \frac{A_i + B_i + D_i}{N}$$

$$PP_{CC,i} = \frac{A_{cc,i} + B_{cc,i} + D_{cc,i}}{A_{cc,i} + B_{cc,i} + C_{cc,i}} = \frac{A_{cc,i} + B_{cc,i}}{N_{cc}}$$

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

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

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

Cumulative distribution function (**CDF**)

The cumulative probability density (Φ_A) can be estimated for the age at vaccination from the subset of persons with a complete follow-up. The cumulative distribution function represents the probability to be vaccinated by a certain age. The increase between week $i - 1$ ($= \Phi_A(t_{i-1})$) and i ($= \Phi_A(t_i)$) is interpreted as the amount of meaningful follow-up (MFU_i). MFU_i thus equals the probability of vaccination during week i inferred from persons with a complete follow-up. $\Phi_A(t_i)$ represents the total amount of meaningful follow-up at the end of week i , $\Phi_A(t_{i-1})$ represents this value at the start of week i .

$$MFU_i = \Phi_A(t_i) - \Phi_A(t_{i-1}),$$

The meaningful follow-up for week i will be multiplied with the proportion of persons in follow-up at week i to obtain the proportion of meaningful follow-up ($MFU_{proportion,i}$).

$$MFU_{proportion,i} = FU_{proportion,i} * MFU_i$$

To allow for age-specific vaccination coverage estimation the proportion of meaningful follow-up at the end of week i will be normalized;

$$MFU_{proportion.normalized,i} = \frac{\sum_{o \rightarrow i} MFU_{proportion,i}}{\sum_{o \rightarrow i} MFU_i}$$

A weighting is done of the total number of vaccinations at the end of week i by the normalized $MFU_{proportion,i}$.

$$CDF_i = \frac{\sum_{o \rightarrow i} A_i}{MFU_{proportion.normalized,i}}$$

Several examples have been published where these methods have been applied to different type of vaccines^{10 11}

Exposure: charts with periodic number of administered doses

The total number of doses n_{ij} (for dose 1, dose 2 and dose x) given during age week i in cohort j should be calculated from each database. This should be done for each database, and type of vaccine.

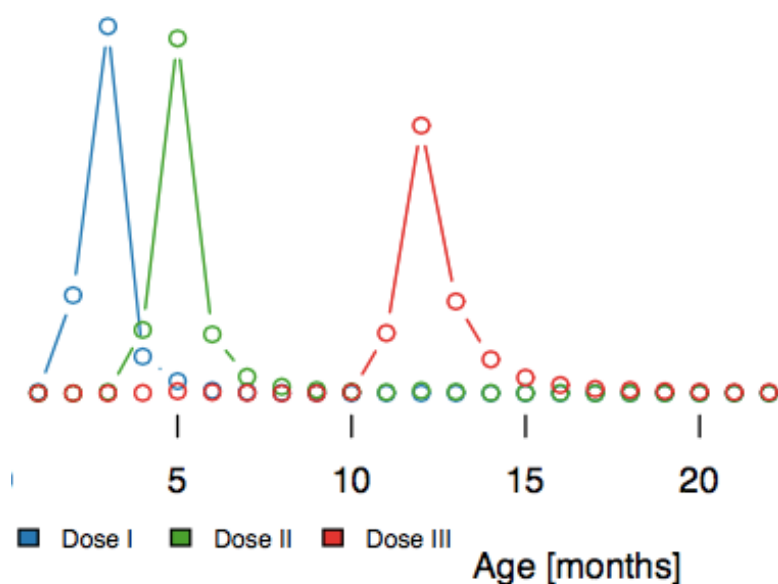


Figure 3: Example of histogram plot with the number of doses by age

Coverage: line plots with vaccination coverage (%) by cohorts over time

For every cohort, vaccination coverage should be calculated by dose over age for each estimation method. The coverage at week i for cohort j will be calculated by dividing the number of vaccinated subjects n_{ij} by the total number of subjects still under follow-up at week i (N_{ij}), expressed as a percentage or by cumulative incidence, and cumulative distribution function

¹⁰ Braeye T, Emborg HD, Llorente-García A, Huerta C, Martín-Merino E, Duarte-Salles T, Danieli G, Tramontan L, Weibel D, McGee C, Villa M, Gini R, Lehtinen M, Titievsky L, Sturkenboom M. 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):3243-3254. doi: 10.1016/j.vaccine.2020.02.082. Epub 2020 Mar 12. PMID: 32171573.

¹¹ Emborg HD, Kahlert J, Braeye T, Bauwens J, Bollaerts K, Danieli G, Duarte-Salles T, Glismann S, Huerta C, de Lusignan S, Martin E, McGee C, Correa A, Tramontan L, Weibel D, Sturkenboom M. ADVANCE system testing: Can coverage of pertussis vaccination be estimated in European countries using electronic healthcare databases: An example. *Vaccine*. 2019 Oct 31;S0264-410X(19)30930-2. doi: 10.1016/j.vaccine.2019.07.039. Epub ahead of print. PMID: 31677953.

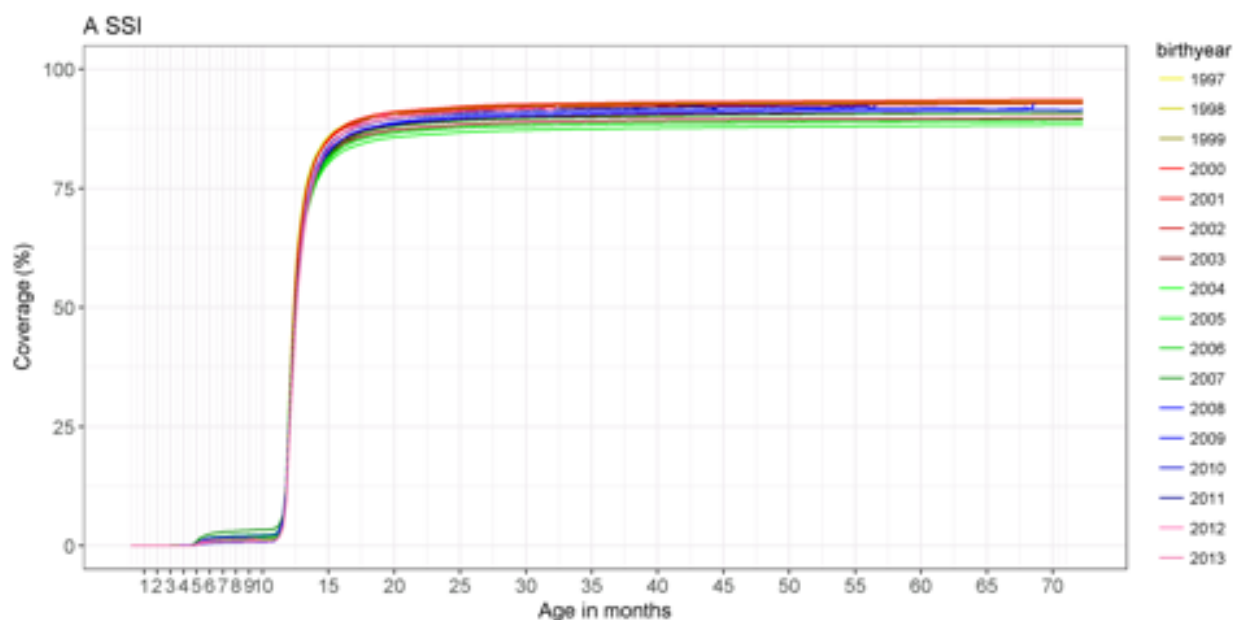


Figure 3: An example of coverage curves from the ADVANCE project

9.6.3 Data Integration

Results should be presented separately for each data source and pooled across data sources. We do not recommend pooling because of differences in vaccination strategies.

9.6.4 Subgroup Analysis

[The study investigator(s) should describe subgroups motivated by the current understanding of the study outcomes in this section.]

If relevant to specific events, the presence of effect modification by relevant variables (age, profession, specific comorbidities) should be assessed using stratification.

9.6.5 Sensitivity Analysis

Sensitivity analysis should focus on the robustness of results to assumptions of the study design and availability of key data elements and should be conducted for the rapid assessment and signal evaluation studies and may include the following:

- If the risk window is not well known, conduct analyses with alternative risk intervals and/or washout periods between the risk and comparison windows
- If exact dates of events are unknown and some are imputed (e.g., if the onset of the event could be prior to date assigned by case validation), conduct analyses lagging the event date.
- If time-varying confounders are not fully available in all data sources, restrict the analysis to sites where data are available on these time-varying confounders and assess at the impact of removing time-varying confounders

9.7 Quality Control

Standard operating procedures or internal process guidance at each research centre should be adhered to for the conduct of the study. These procedures should include internal quality audits, rules for secure and confidential data storage, methods to maintain and archive project documents, quality-control procedures for programming, and standards for writing analysis plans.

This section should describe the study-specific process for programming quality control (e.g., independent programming and/or review of summary output and programming logs by a second programmer), and procedures for data storage, archiving, and backup at each study centre. Also described should be processes for review and quality control of study documentation and reporting of pooled results across research centres.

Note to future investigators: The pandemic has led to changes in healthcare utilization and provision which are likely to extend into the vaccine roll-out period. This may be reflected in observational data as an excess of code counts for a subset of diseases and/or their proxies in the pandemic period, or as a deficit for others.

9.8 Limitations of the Research Methods

The proposed study design is subject to limitations due to both the study design and secondary use of health care data/registries.

Data-related limitations include dependency on the accuracy of codes and algorithms to identify at risk conditions, professions or living circumstances.

Exposure ascertainment may be based on pharmacy dispensing records, general practice records, immunization registers, medical records, or other electronic data sources. The ability to identify specific COVID-19 vaccines and dates of vaccination are currently unknown as it is not clear how the vaccines will be rolled out and what level of detail will be recorded. ACCESS promotes the recording and identification of vaccine brands and batch numbers/ lot numbers. It is likely that subjects vaccinated outside of the healthcare system may not have a record of their vaccination. If brands cannot be distinguished, there may be misclassification of exposure which is of essence due to the differences in platforms and adjuvants.

Unvaccinated subjects may become exposed to COVID-19 vaccine at any time over the course of study, censoring in the follow-up time may create an imbalance in the observation periods between vaccinated and unvaccinated groups, which is not random.

9.9 Other Aspects

This section, which is optional, should contain information on any other aspect of the research method not covered by previous sections, such as scientific advisory board or endpoint adjudication committees.

10 Protection of Human Subjects

The proposed studies are non-interventional studies re-using health care data. All data collected in the study will be de-identified with no breach of confidentiality with regard to personal identifiers or health information. Each data access provider should apply for an independent ethics committee review according to local regulations and the local DPIA should be informed. Data protection and privacy regulations (GDPR) should be respected in collecting, forwarding, processing, and storing data from study participants.

11 Management and Reporting of Adverse Events /Adverse Reactions

Aggregate analysis of database studies can identify an unexpected increase in risk associated with a particular exposure. Such studies may be reportable as study reports, but typically do not require reporting of individual cases. Moreover, access to automated databases does not confer a special obligation to assess and/or report any individual events contained in the databases. Formal studies conducted using these databases should adhere to these guidelines.

For non-interventional study designs that are based on secondary use of data, such as studies based on medical chart reviews or electronic health care records, systematic reviews or meta-analyses, reporting of adverse events/adverse drug reactions is not required. Reports of adverse events/adverse drug reactions should only be summarised in the study report, where applicable (EMA, 2017)

According to the EMA Guideline on Good Pharmacovigilance Practices (GVP), Module VI – Management and Reporting of Adverse Reactions to Medicinal Products (EMA, 2017),

“All adverse events/reactions collected as part of [non-interventional post-authorisation studies with a design based on secondary use of data], the submission of suspected adverse reactions in the form of [individual case safety reports] is not required. All adverse events/reactions collected for the study should be recorded and summarised in the interim safety analysis and in the final study report.”

Module VIII – Post-Authorisation Safety Studies, echoes this approach (EMA, 2020). The new legislation further states that for certain study designs such as retrospective cohort studies, particularly those involving electronic health care records, it may not be feasible to make a causality assessment at the individual case level.

12 Plans for Disseminating and Communicating Study Results

In its Guidelines for GPP, ISPE contends that “there is an ethical obligation to disseminate findings of potential scientific or public health importance” (ISPE, 2015); for example, results pertaining to the safety of a marketed medication. “...the marketing authorisation holder should communicate to the Agency and the competent authorities of the Member States in which the product is authorised the final manuscript of the article within two weeks after first acceptance for publication.”

Protocols should be registered at the EU PAS register and comply with ENCePP or ADVANCE code of conducts. According to both codes of conducts

Study results will be published following guidelines, including those for authorship, established by the International Committee of Medical Journal Editors (ICMJE, 2019). When reporting results of this study, the appropriate Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist will be followed (STROBE, 2015), and recommendations on reproducible reporting of electronic health care data base studies should be followed (Wang, 2017)

Communication via appropriate scientific and regulatory venues should be considered.

[To be completed or modified by study investigator(s), as needed.]

13 Other Good Research Practice

This study will adhere to the *Guidelines for GPP* and has been designed in line with the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCEPP) *Guide on Methodological Standards in Pharmacoepidemiology* (ENCEPP, 2018). The *ENCEPP Checklist for Study Protocols* (ENCEPP, 2018) will be completed (see Annex 2).

The study is a post-authorisation study of vaccine coverage and will comply with the definition of the non-interventional (observational) study referred to in the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use tripartite guideline *Pharmacovigilance Planning E2E* (ICH, 2019) and provided in the EMA *Guideline on Good Pharmacovigilance Practices (GVP) Module VIII: Post-Authorisation Safety Studies* (EMA, 2020), and with the 2012 European Union pharmacovigilance legislation, adopted June 19, 2012 (European Commission, 2019). The study will comply with the study reporting requirements specified in Module VIII section VIII.B.6.3.1. “Progress reports” and VIII.B.6.3.2. “Final Study Report” of the *Guideline of Good Pharmacovigilance Practices* (EMA, 2020).

The study will be registered in the European Union Post-Authorisation Study Register (ENCEPP, 2019) before the study implementation commences.

The research team and study sponsor should adhere to the general principles of transparency and independence in the ENCePP Code of Conduct (ENCEPP, 2020) or the ADVANCE code of conduct (Kurz, 2017)

[If desired by the study investigators, the following may be included] The research team will apply for the ENCePP Study Seal (ENCEPP, 2018).

[To be completed or modified by the study investigator(s), as needed. Country-specific study registration requirements may be discussed here, where required.]

14 References

1. EMA: European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). 22 November 2017. Available at: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf. Accessed 16 June 2020.
2. EMA: European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module VIII – Post-authorization safety studies (EMA/813938/2011 Rev 3). 13 October 2017. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf. Accessed 17 January 2020.
3. ENCePP. The ENCePP code of conduct for scientific independence and transparency in the conduct of pharmacoepidemiological and pharmacovigilance studies (Revision 4). 15 March 2018. Available at: http://www.encepp.eu/code_of_conduct/. Accessed 17 January 2020.
4. ENCePP: European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. ENCePP checklist for study protocols (revision 4). 15 October 2018. Available at: http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml. Accessed 15 January 2019.
5. ENCePP: European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Guide on methodological standards in pharmacoepidemiology (EMA/95098/2010 Rev. 7). July 2018. Available at: http://www.encepp.eu/standards_and_guidances/. Accessed 15 January 2019.
6. ENCePP: European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. The ENCePP seal. 20 December 2018. Available at: http://www.encepp.eu/encepp_studies/index.shtml. Accessed 15 January 2019.
7. ENCePP: European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. The European Union Electronic Register of Post-Authorisation Studies (EU PAS Register). 20 December 2018. Available at: http://www.encepp.eu/encepp_studies/indexRegister.shtml. Accessed 15 January 2019.
8. European Commission. Commission implementing regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council. 20 June 2012. Available at: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF>. Accessed 17 January 2019.
9. Gini, Rosa, et al. "Different Strategies to Execute Multi-Database Studies for Medicines Surveillance in Real-World Setting: A Reflection on the European Model." *Clinical Pharmacology & Therapeutics* (2020).
10. ICMJE: International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. December 2018. Available at: <http://www.icmje.org/recommendations/>. Accessed 17 January 2019.
11. ISPE: International Society for Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practices (GPP). Revision 3. June 2015. Available at:

- <https://www.pharmacoepi.org/resources/policies/guidelines-08027/>. Accessed 17 January 2020.
12. Kurz, Xavier, et al. "The ADVANCE Code of Conduct for collaborative vaccine studies." *Vaccine* 35.15 (2017): 1844-1855.
 13. Li, Xiaojuan, et al. "Validity of Privacy-Protecting Analytical Methods That Use Only Aggregate-Level Information to Conduct Multivariable-Adjusted Analysis in Distributed Data Networks." *American journal of epidemiology* 188.4 (2018): 709-723.
 14. STROBE: Strengthening the Reporting of Observational Studies in Epidemiology. Checklists. 2007. Available at: <http://www.strobe-statement.org/index.php?id=available-checklists>. Accessed 05 January 2015.
 15. Wang, Shirley V., et al. "Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1. 0." *Value in health* 20.8 (2017): 1009-1022.

Annex 1. List of Stand-Alone Documents

Number	Document reference number	Date	Title
1	Annex 2	<<MM.DD.YYYY>>	ENCePP checklist for protocols

Annex 2. ENCePP Checklist for Study Protocols

The ENCePP Checklist for Study Protocols can be accessed and downloaded using the following link:
[http://www.encepp.eu/standards and guidances/checkListProtocols.shtml](http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml)