www.advance-vaccines.eu

Accelerated Development of VAccine beNefit-risk Collaboration in Europe

IMI JU Grant Agreement nº115557

Proof of Concept Study Protocol

Exposure and coverage to routine schedule vaccines in different EU countries

WP5 – Proof-of-concept of a framework
to perform vaccine benefit-risk monitoring

Disclaimer: *The study described in this protocol are conducted as part of the IMI ADVANCE project with the aim to test methodological aspects of the design, conduct and reporting of studies for vaccine benefit-risk monitoring activities.*

The protocol presented herein relates solely to the testing of these methodologies and is not intended to inform regulatory or clinical decisions on the benefits and risks of the exposures under investigation. Therefore any use of information from these studies should carry over this warning and be used accordingly.

V 1.0

Title	Exposure and coverage to routine schedule vaccines in different EU countries
Medicinal product	All types of vaccines that were available during the study period and used in the participating study population in the participating countries, as part of the routine childhood or elderly vaccination program. These include vaccines against the following vaccine preventable diseases Tuberculosis Rotavirus Diphteria Tetanus Pertussis Poliomyelitis Haemophilus influenzae type B Hepatitis B Pneumococcal disease Meningococcal disease Measles Mumps Rubella Varicella Human papillomavirus Influenza Herpes Zoster
Product reference	Any vaccine protecting against any of the above listed vaccine preventable diseases
Research question and objectives	The primary objective is to determine coverage to vaccines that are part of the routine immunization schedule. The secondary objective is to compare the estimates with external benchmarks
Countries of study	The project will include all data sources that are accessible to partners of the ADVANCE consortium, and have provided their willingness to participate. This includes data sources from: United Kingdom, Italy, Denmark, Spain and the Netherlands

1. TABLE OF CONTENTS

2. DEFINITIONS

- Participants of the ADVANCE Consortium are referred to herein according to the following codes:
  - **P95.** P95 (Belgium)
  - **ARS:** Agenzia, Regionale di Sanita, Toscana (Italy)
  - **AEMPS.** Agencia Española de Medicamentos y Productos Sanitarios (Spain)
  - **ATSVP. Agenzia di Tutela della Salute della Val Padana** (Italy)
  - **AUH.** Aarhus Universitetshospital (Denmark)
  - **J&J.** Janssen Vaccines - Prevention B.V. (The Netherlands)
  - **ECDC. E**uropean Centre for Disease Prevention and Control (Sweden)
  - **EMA.** European Medicines Agency (United Kingdom)
  - **EMC.** Erasmus Universitair Medisch Centrum Rotterdam (Netherlands)
  - **GSK.** GlaxoSmithKline Biologicals, S.A. (Belgium)
  - **IDIAP-Jordi Gol,** Jordi Gol Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Barcelona (Spain)
  - **KI.** Karolinska Institutet (Sweden)
  - **LSHTM.** London School of Hygiene and Tropical Medicine (United Kingdom)
  - **MHRA.** Medicines and Healthcare products Regulatory Agency (United Kingdom)
  - **NOVARTIS/Seqirus*.** Novartis Pharma AG (Switzerland)
  - **OU.** The Open University (United Kingdom**)**
  - **PEDIANET**. Società Servizi Telematici SRL (Italy)
  - **PFIZER**. Pfizer Limited (United Kingdom)
  - **RCGP_RSC.** Royal College of General Practitioners Research and Surveillance Centre (United Kingdom)
  - **RIVM.** National Institute for Public Health and the Environment (Netherlands)
  - **SP.** Sanofi Pasteur (France)
  - **MSD** Merck Sharp & Dohme Corp (USA)
  - **SSI.** Statens Serum Institut (Denmark)
  - **SURREY.** The University of Surrey (United Kingdom)
  - **SYNAPSE.** Synapse Research Management Partners, S.L. (Spain)
  - **TAKEDA.** Takeda Pharmaceuticals International GmbH (Switzerland)
  - **UNIBAS.** Universitaet Basel (Switzerland)
  - **UTA.** Tampereen Yliopisto (Finland)
  - **Sciensano** Belgian Scientific Institute of Public Health (Belgium)

  ** Effective 9 November 2015, bioCSL, the vaccine and pharmaceutical business of CSL, acquired the influenza vaccines business of Novartis, to create Seqirus, a CSL company. Seqirus and Novartis operate at interim under the Sale and Purchase Agreement governing the sale to CSL as well as the relevant TSAs and TDSA.*

  Associate partners are referred to herein according to the following codes:
  - **AIFA**: Italian Medicines Agency (Italy)
  - **ANSM**: French National Agency for Medicines and Health Products Safety (France)
  - **BCF** Brighton Collaboration Foundation (Switzerland)

- **EOF** Helenic Medicines Agency, National Organisation for Medicines (Greece)
- **FISABIO** Foundation for the Promotion of Health and Biomedical Research(Spain)
- **HCDCP** Hellenic Centre for Disease Control and Prevention (Greece)
- **ICL** Imperial College London (UK)
- **IMB** Irish Medicines Board (Ireland)
- **IRD** Institut de Recherche et Développement (France)
- **NCE** National Center for Epidemiology (Hungary)
- **NSPH** Hellenic National School of Public Health (Greece)
- **PHE** Public Health England (UK)
- **THL** National Institute for Health and Welfare (Finland)
- **UOA** University of Athens (Greece)
- **UNIME** University of Messina (Italy)
- **UMCU** University Medical Center Utrecht (Netherlands)
- VACCINE.GRID foundation (Switzerland)
- **WKT** State Medicines Control Agency (Lituania)
- **WUM** Polish Medicines Agency (Poland)

- **Aggregated Data:** The cumulative or summary information that does not specifically identify any particular person[1]
- **Anonymised Data:** Data that cannot be traced back to the individual patient
- **Harmonised Data:** The harmonized data follow a consensus and are formatted in the same way
- **Original Data:** Data, as maintained by the Data Source or any organization which collects the data, before inclusion in the platform
- **Personal data:** any information relating to an identified or identifiable natural person (data subject); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity
- **Identified or identifiable natural person:** means anyone who "can be identified, directly or indirectly, in particular by reference to an identification number or by one or more factors specific to his physical, physiological, mental, economic, cultural, or social identity."
- **Processing of personal data:** means "any operation or set of operations . . . performed upon personal data," automatically or otherwise this definition is wide open, because it includes "collection, recording, organization, storage . . . retrieval . . . use, disclosure by transmission," and "dissemination." By expressly including "storage" in the definition of "processing," the mere act of holding personal data is, under EU law, a regulated activity.
- **Data "controller":** the "'controller' shall mean the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data
- **Data "processor":** is anyone who processes personal data for a controller
- **Third party:** is anyone who processes data under "the direct authority" of a controller or processor

[1] https://help.blackboard.com/en-us/Learn/9.1_SP_14/Administrator/050_Security/010_Privacy/000_US_Privacy_Definitions_and_Regulations

- **Database fingerprinting:** characterization of the actual data content of a database
- **Study:** de novo information generation following a specific question

3. RESPONSIBLE PARTIES

3.1 Main Author(s) of the Protocol & document history

Name	***Institution***	***Role***	***Contribution***
Miriam Sturkenboom	P-95	Investigator	Drafting protocol
Vincent Bauchau, Hanne-Dorthe Emborg, Piotr Kramarz, Johnny Kahlert, John Weil, Tin Tin Htar Myint, Patricia Saddier, Alena Khromava, Lieke van der Aa, Simon de Lusignan, Lina Titievsky	GSK, SSI, ECDC, AUH, Takeda, Pfizer, Merck, Sanofi Pasteur, Sciensano, RCGP RSC	SC and study team members	Comments
Toon Braeye	Sciensano	Study team, statistician	Comments v0.2/0.3
Miriam Sturkenboom	P-95	Investigator	Addressing comments Toon & adding in Praeventis database, taking out Finnish cohort on request of Braeye as it is a selected cohort
Patricia Saddier	MSD	SC member	Comments v0.4 request for more details on analysis
Miriam Sturkenbom	P-95	Investigator	Addition of analysis details from the statistical analysis plan and example graphics from prior study
Miriam Sturkenboom, Lina Titievsky, Lieke van der Aa, Toon Braeye		WP5 leaders	Addressing comments from in house reviews/comments of SC and companies

3.2 Principal Investigator

Miriam Sturkenboom, PhD, PharmD, P-95

3.3 Study Team

Lina Titievsky, MD, PhD, Pfizer

Toon Braeye, MD, statistician, Sciensano
Hanne-Dorthe Emborg, D.V.M, PhD, epidemiologist, SSI
Johnny Kahlert, PhD, statistician, Aarhus
Chris McGee, analyst RCGP RSC
Simon de Lusignan, RCGP RSC, quality control RCGP RSC
Mark Joy, RCGP-RSC quality control RCGP RSC
Rosa Gini, PhD, ARS
Lara Tramontan, PhD, PEDIANET
Giorgia Danieli, analyst, PEDIANET
Elisa Martin, PhD BPharm, epidemiologist, BIFAP
Talita Duarte Salles, PhD, epidemiologist, SIDIAP
Marco Villa, PhD, statistician, ATSVP
Lieke van der Aa, PhD, project manager, Sciensano
Matti Lehtinen, PhD, physician, Finnish cohorts, University Tampere
Hester de Melker, MD, PhD, epidemiologist RIVM, The Netherlands
Susan Hahne, PhD, epidemiologist RIVM (Praeventis lead contact)

4. Abstract

Date of Protocol Abstract: February 1, 2019
Exposure and coverage to routine schedule vaccines in different EU countries
Rationale and Background: To increase readiness of European data sources to conduct multi-site vaccine coverage, benefit and risk studies
Research Question and Objective: The primary objective is to determine exposure in the population, and coverage of routine vaccinations in children and elderly The secondary objective is to coverage estimates with external benchmarks
Study Design: The main study design is a retrospective dynamic cohort study
Population: The source population comprises of all subjects registered with one of the ADVANCE participating databases, which can link vaccination and population data and have provided their willingness to participate. The study population comprises all persons who are registered during the study period that starts January 1, 2000 and ends December 31, 2018.
Outcome Parameters: Vaccine exposure and vaccine coverage The following outcome parameters will be assessed: Number of persons in the population by database Number of persons vaccinated, by vaccine, by dose by database Number of doses, brands and vacine types administered in each database Vaccine coverage by vaccine, per dose, by birthyear and by database
Data Sources:

Electronic health care databases (vaccine registries, record linkage, surveillance and GP based databases) currently available in the ADVANCE consortium and eligible are located in Denmark, Spain, Italy, UK and Netherlands Informative data sources Vaccine schedules: European Centre for Disease Prevention and Control (ECDC) vaccine schedules in Europe and national statistics on vaccination coverage Benchmark data: WHO estimates for following vaccines/doses in a country: BCG, DTP1, DTP3,HepB_BD, HepB3, Hib3, IPV1, MCV1, MCV2, PCV3, Pol3, RCV1, Rota, http://apps.who.int/immunization_monitoring/globalsummary/timeseries/tswucoveragedtp3.html Published literature for HPV (Bruni, Lancet Global Health) and influenza (VENICE)
Study Size: Total population in eligible ADVANCE databases for exposure. Coverage estimates by birth cohorts 2000-2017 for childhood vaccination and birthyears 1930-1960 for vaccines targeted to elderly
Data Analysis: Frequencies and distributions are measured by general descriptive statistics. Exposures will be the number of vaccinations administered. Coverage rates are presented by three different estimators. These included; period prevalence crude and corrected for follow-up and cumulative incidence
Informed Consent and Ethical Approval: This study will be conducted on the basis of secondary use of electronic healthcare records. Each database access provider will apply local governance and privacy rules prior to aggregating and sharing anonymized data.
Milestones: August 2018 Outline to Steering Committee September Discussion in General Assembly December Protocol submission for in house clearance December –January Conversion of data into common data model (feasibility) February 2019 Data transformation & sharing of results 31, March 2019 Report delivered

Amendments and Updates

Protocol amendments following IRB approval:

Table 1: Overview of Protocol Amendments and Updates

Number	***Date (DDMMMYY)***	***Section of the study protocol***	***Amendment or update***	***Reason***
1				
2				
….				

MILESTONES

Table 2: Overview of Study Milestones

August 2018 Outline to SC

September	Discussion in GAM
December	Protocol submission for in house clearance
December –January	Conversion of data into common data model locally (feasibility)
February 2019	Data transformation & sharing of results
31, March 2019	Report delivered

5. Rationale and Background

The ADVANCE project emerged from the need to build an integrated and sustainable framework for the continuous monitoring of the benefit/risk of vaccines in Europe.

The ADVANCE vision is to deliver "**Best evidence at** the **right time** to support **decision-making** on **vaccination** in Europe"**,** and the mission is to establish a prototype of a sustainable and compelling system that rapidly provides best available scientific evidence on vaccination benefits and risks post-licensure for well informed decisions.

A first set of proof of concept studies were conducted, with 4 different studies on pertussis vaccine coverage, vaccine benefits, vaccine risk and vaccine benefit/risk of pertussis containing vaccines (www.advance-vaccines.eu) The experiences and lessons learned were described in a white paper. One of the recommendations was to look at brand specific data and more vaccines.
The second proof of concept (POC-2) plan was developed to further increase readiness of databases to conduct vaccine studies by estimating coverage to routine vaccine schedules and comparing them with external benchmark data.

6. RESEARCH QUESTION AND OBJECTIVES

6.1 Aim:

To increase readiness of European data sources to conduct multi-site vaccine coverage, benefit and risk studies

6.2 Primary Objective

The objective is to determine exposure and coverage to vaccines that are part of the routine immunization schedule

6.3 Secondary objective

To compare the coverage estimates (cumulative distribution function) with published national and WHO benchmarks, to assess whether data sources are fit for purpose for specific studies.

7. RESEARCH METHODS

7.1 Study Design

A retrospective dynamic population-based cohort analysis, during the study period 2000-2018

7.2 Setting

The analysis will be conducted in multiple population based healthcare databases in various European countries

7.3 Databases/Data Sources

The POC-2 study will be conducted on data in electronic healthcare databases that reside with partners and associate partners of the ADVANCE consortium who have expressed interest to participate. A short description of each of the data sources is listed below.

The project will include all data sources that are accessible to partners of the ADVANCE consortium, and have provided their willingness to participate. Data sources can roughly be divided between record linkage data sources where regional/national hospitalization discharge data are linked to vaccination registries and population files (e.g. SSI, AUH, ASL Cremona, ARS, Praeventis) and general practice databases (RCGP, PEDIANET, SIDIAP, BIFAP), SIDIAP and PEDIANET can also be linked to other data sources.

SSI, Denmark
The Danish Civil and Health Registration System database is created ad hoc by linkage between the Danish civil registration system, the vaccination registry, the patient registry plus other relevant databases (e.g., disease surveillance, medications, microbiology, and so on).
The data are kept at Research Services at The Danish Health Data Authority from where access can be granted provided that the analysis has been reported and approved by the Danish Data Protection Agency. The vaccination registry includes personal ID for linkage. Also the database is not a single database but data are assembled and merged for specific studies according to the relevant Scientific questions. All diagnoses are coded using the ICD-10 (Danish modification)

Aarhus, Denmark
The Aarhus University Prescription database comprises clinical and prescription data on the population of former North-Jutland, Aarhus, Ringkøbing and Viborg counties, which since 2007 are called the Central Denmark Region and the North Denmark Region. This population covers a total of 1.8 million inhabitants and is representative of the population of Denmark. Data available on these subjects comprise their eligibility, dispensing data, hospitalizations and procedures and the population can also be linked to other National Danish registries. Dispensing data comprise the filled prescriptions for all ambulatory patients and contains information on name of the drug, ATC code, package identifier (strength and route of administration), and the date of refill. These data can be linked to the national registry of patients that comprises information on admissions to Danish somatic hospitals, emergency rooms and outpatient clinics, diagnosis codes and procedures are registered. These databases have been used in numerous studies and are proven valid for pharmaco-epidemiological research. All diagnoses are coded using the ICD-10 (Danish modification).

BIFAP, Spain
BIFAP (Base de Datos para la Investigación Farmacoepidemiológica en Atencion Primaria) is a computerised database of medical records of Primary Care operated by the Spanish Medicines Agency (AEMPS) as a non-profit research project. In the moment AVANCE started, the database included data from 5,714 physicians (including general practitioners and paediatricians) from 9 different regions in Spain

between 2003-2014. These regions collaborate with BIFAP and send their data to BIFAP every year. BIFAP database includes clinical and prescription data from around 7,5 million patients covering around 16.4% of the Spanish population. Diagnoses are coded using ICPC and free text. The database is restricted to research for regulatory purposes and to independent research by the academia and healthcare professionals of the national health system. In the BIFAP database vaccines provided by the pediatrician/GP are recorded using local codes. A description for the BIFAP databases is included in the ENCePP register: http://www.encepp.eu/encepp/viewResource.htm?id=21501

SIDIAP database, Spain

GPs play an essential role in the public health care system of Spain, as they are responsible for primary health care, long-term prescriptions and specialist and hospital referrals. The Spanish public health care system covers more than 98% of the population. The Information System for Research in Primary Care (SIDIAP; www.sidiap.org) database comprises of anonymized electronic medical records of the population attended by a health professional in a primary care centre of the Catalan Health Institute, the main health provider in Catalonia (North-East Spain). It covers a population of more than 5.8 million people (about 80% of the total of 7.5 million population of Catalonia) from 287 primary care practices with 3,414 participating GPs. The SIDIAP data comprises the clinical and referral events registered by primary care health professionals (GPs and nurses) in electronic medical records, comprehensive demographic information, prescription and corresponding pharmacy invoicing data, specialist referrals, and primary care laboratory test results. Hospital admissions and their major outcomes can be identified for a number of practices, covering a total 1.7 million active patients. Health professionals gather this information using ICD-10 codes (primary care records) and ICD-9 (hospital admissions), and structured forms designed for the collection of variables relevant for primary care clinical management, such as nationality, sex, age, height, weight, body mass index, tobacco and alcohol use, blood pressure measurements, blood and urine test results. Encoding personal and clinic identifiers ensures the confidentiality of the information in the SIDIAP database. Recent reports have shown the SIDIAP data to be useful for epidemiological research (Garcia-Gil et al 2011).

As this is a primary care database, information on specialist prescribing, drug dispensing in hospital setting and actual drug intake is missing. The database is listed in the ENCePP catalogue: http://www.encepp.eu/encepp/viewResource.htm?id=4646

ATS della Val Padana (ATSVP), Italy

As the effect of the Regional Law n. 23/2015, on 01/01/2016 ASL Cremona and ASL Mantova ceased to exist and merged in a new institution called ATS (Agenzia di Tutela della Salute) della Val Padana. Such Local Health Agency (ATS) is the institution in charge of the health of the citizens living in the provinces of Cremona and Mantova and is responsible for the governance and control of all the health-related services (prevention, treatment, residential care, etc.). ATSVP is a record linkage database. It contains all the mortality data (with cause of death), hospitalizations (with diagnosis), outpatient visits, drug prescriptions of the citizens. Moreover, it contains the registry of all the vaccinations administered by (or notified to) the Local Health Agency and the registry of infectious diseases to be notified by law. All diagnoses are coded using the ICD-9. In its vaccine register, all routine child vaccines are captured including brand and dose.

Pedianet, Italy

Pedianet is a primary care paediatric medical record database including data collected by primary care paediatricians (FPs - Family Practitioners) during their routine daily practice since 2003. This system is based on the transmission of specific data (determined by individual studies) from computerised clinical files, which the paediatricians in the network fill out during their daily professional activities. Informed

consent is required from the parents. Such data are collected anonymously by a central server in Padua, where the data are validated and elaborated. Pedianet allows linkage capabilities to other databases such as the Veneto Region vaccine register via patient code. Each patient is identified by the system from an ID-patient and ID paediatric, informed consent is required for linkage. All diagnoses are coded using ICD-9 and/or captured in free text. In the vaccine register of the Veneto Region, all routine child vaccines are captured including brand and dose. Pedianet is registered in ENCePP catalogue http://www.encepp.eu/encepp/viewResource.htm?id=20131.

Agenzia regionale di Sanita' (ARS) database, Italy

ARS is a record linkage database. It contains a copy of the healthcare administrative databases of the Regional Healthcare System of Tuscany, linked to some registries. It contains all the mortality data (with cause of death), hospitalizations (with diagnosis), access to emergency care (with diagnosis), exemptions form copayment (with diagnosis), pathology registry (with diagnosis), access to community mental health care (with diagnosis), outpatient visits (without diagnosis), and drug dispensings (from primary care or specialist prescribing) of the inhabitants of the Tuscany region. Moreover, since recently it contains the registry of the vaccinations administered by the personnel of the Regional Healthcare System . All diagnoses and procedures are coded using the ICD-9CM, except the pathology registry with is coded in SNOMED and has free text diagnostic fields. In the vaccine register, all routine child vaccines are captured including brand and dose. The database description is recorded in the ENCePP catalogue http://www.encepp.eu/encepp/viewResource.htm?id=24417

Praeventis, The Netherlands

To monitor the vaccination coverage in the Netherlands, an electronic national immunisation register called 'Præventis' was implemented in 2005. Præventis has a link with the population register and can produce letters of invitation for the national immunization program, register and validate administered vaccinations. The database is used to monitor the vaccination process, produce reminder letters, control the stock of vaccines and provides information used for paying the fees to the different executive organisations involved. Præventis provides a crucial tool for the evaluation of the NIP by producing (sub)national vaccination coverage estimates with high accuracy and allowing additional research: identifying populations at high risk for low coverage based on existing data, conducting specific studies where individuals included in the immunisation register are approached for further research, using vaccination coverage data for the interpretation of (sero)surveillance data, and linking the immunisation register with disease registers to address vaccine safety or vaccine effectiveness. The ability to combine Præventis data with data from other databases or disease registers and the ability to approach individuals with additional research questions offers opportunities to identify areas of priority for improving the Dutch NIP (van Lier et al.)[2]

RCGP RSC, UK

RCGP RSC is a primary care sentinel network and has been set up for monitoring influenza and respiratory disease surveillance and has been supported by the Department of Health (DH) since 1967[3]. The RCGP Research and Surveillance Center has grown in terms of size, scope, integration of microbiological testing, data-linkage[4].

[2] Præventis, the immunisation register of the Netherlands: a tool to evaluate the National Immunisation Programme van Lier, A and Oomen, P and de Hoogh, P and Drijfhout, I and Elsinghorst, B and Kemmeren, J and Conyn-van Spaendonck, M and de Melker, H, Eurosurveillance, 17, 20153 (2012), https://doi.org/10.2807/ese.17.17.20153-en

[3] de Lusignan S, Correa A, Smith GE, Yonova I, Pebody R, Ferreira F, Elliot AJ, Fleming D. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. Br J Gen Pract. 2017 Oct;67(663):440-441. doi: 10.3399/bjgp17X692645.

Between 1967 and 1993 the registered population within the RCGP RSC grew to 200,000 and it comprised 40 practices. By its 30th anniversary the population had risen to 570,000, with electronic links to practices and laboratory data. The network exceeded 1 million in 2015, and has now expanded to 417 practices and 4.2 million patients.

The network has been representative of the national population. Initially the denominator was based on the practices own age-sex register, more recently this has become electronically validated against national extracted denominators. We provide patient level socio-economic status and ethnicity, where recorded. The disease data was recorded in practice based diagnostic indexes. Computer recording of diagnostic data in the 1990's facilitated expansion of analytical possibilities including trend analysis of secular change in the incidence of common diseases and of chronic conditions. In 2015 the database was extended to collect all Read coded data on an individual patient basis but pseudonymised.

In collaboration with the Public Health Laboratory Service, the network began collecting respiratory virology samples from patients presenting with acute influenza-like-illness in the early 1990s to enhance influenza surveillance, allowing the onset of the season to be detected, the circulating strains characterised and the intensity monitored. It has periodically been involved in pilots of chlamydia, shingles, and gastroenteritis infection. It has also carried out detailed data collections alongside several of the 10-yearly national census. Its system of pseudonymisation allows the linkage of RCGP RSC data to other health data, without revealing patients' identities.

Variables include a household key, able to measure household transmission, [3] urban-rural classification,[4] and ethnicity and socioeconomic status.[5] A dashboard technology enables regular feedback to practices about data quality and can be used for quality improvement and trials.[6]

7.4 Source Population

The source population in each of the databases will be the population (all ages) registered for at least one day during the study period (2000-2018). For coverage estimations the population will be restricted to birthyears 2000-2017 (childhood vaccines) and 1930-1960 (vaccines for elderly)

7.5 Study Population Selection

The study population comprises all persons who are registered for at least one day during the study period that starts January 1, 2000 and ends December 31, 2018. Follow-up starts at the first date the person is

[3] de Lusignan S, Konstantara E, Joy M, Sherlock J, Hoang U, Coyle R, Ferreira F, Jones S, O'Brien SJ. Incidence of household transmission of acute gastroenteritis (AGE) in a primary care sentinel network (1992-2017): cross-sectional and retrospective cohort study protocol. BMJ Open. 2018 Aug 23;8(8):e022524. doi: 10.1136/bmjopen-2018-022524.

[4] de Lusignan S, McGee C, Webb R, Joy M, Byford R, Yonova I, Hriskova M, Matos Ferreira F, Elliot AJ, Smith G, Rafi I. Conurbation, Urban, and Rural Living as Determinants of Allergies and Infectious Diseases: Royal College of General Practitioners Research and Surveillance Centre Annual Report 2016-2017. JMIR Public Health Surveill. 2018 Nov 26;4(4):e11354. doi: 10.2196/11354.

[5] de Lusignan S, Correa A, Pathirannehelage S, Byford R, Yonova I, Elliot AJ, Lamagni T, Amirthalingam G, Pebody R, Smith G, Jones S, Rafi I. RCGP Research and Surveillance Centre Annual Report 2014-2015: disparities in presentations to primary care. Br J Gen Pract. 2017 Jan;67(654):e29-e40. doi: 10.3399/bjgp16X688573.

[6] Pathirannehelage S, Kumarapeli P, Byford R, Yonova I, Ferreira F, de Lusignan S. Uptake of a Dashboard Designed to Give Realtime Feedback to a Sentinel Network About Key Data Required for Influenza Vaccine Effectiveness Studies. Stud Health Technol Inform. 2018;247:161-165.

registered in the databases and ends at the end of the study period, death, moving out of database registration, whichever date is earliest.

7.6 Study Period

The study period is from 01 January 2000 to 31 December 2018 and be dependent on availability of data from specific databases.

7.7 Variables

All variables will be obtained from the participating electronic health care databases. These will be transformed into a pre-defined ADVANCE common data model (CDM) to allow for the running of common scripts across all the databases. Population data of interest will be the coded patient identifier, the date of birth, the start and end of follow-up and the gender. This will be transferred into the common data model input files.

7.7.1 Exposure of Interest, Operationalization and Validation

Exposure of interest

The exposure of interest in this project will be all types of vaccines that were available during the study period and used in the participating study population in the participating countries, as part of the routine childhood or elderly vaccination program. Routine schedules across the different countries as obtained from the ECDC Vaccinescheduler comprise currently. Data on vaccinations will be extracted from 2000 onwards.

Table 4: Current routine immunization schedule in participating countries (from ECDC vaccine scheduler)

Type	Italy	Spain	UK	Denmark	Netherlands
Tuberculosis			BCG birth (at risk)		
Rotavirus	2-7 months (2 or 3 doses)		2,3 months		
Diphtheria	3,5,11 months, 6 years, 12-18 yrs, pregnancy (Tdap)	2,4, 11 months, 6 years, 14 years, >65 yrs	2,3,4 months, 3, 14 years	3,5,12 months, 5 yrs.	1-2, 3,4,11 months, 4 years, 9 years
Tetanus	3,5,11 months, 6 years, 12-18 yrs, pregnancy (Tdap)	2,4, 11 months, 6 years, 14 years, >65 yrs	2,3,4 months 3,14 years	3,5,12 months, 5 yrs.	1-2, 3,4,11 months, 4 years, 9 years
Pertussis	3,5,11 months, 6 years, 12-18 yrs, pregnancy (Tdap)	2,4, 11 months, 6 years	2,3,4 months 3 years,15-45 pregnancy	3,5,12 months, 5 yrs.	1-2, 3,4,11 months, 4 years
Poliomyelitis	3,5,11 months, 6 years, 12-18 yrs.	2,4, 11 months, 6 years	2,3,4 months 3,14 years	3,5,12 months, 5 yrs.	1-2, 3,4,11 months, 4 years, 9 years
Haemophilus influenzae type B	3,5,11 months	2,4, 11 months, 6 years	2,3,4, 12 months	3,5,12 months	1-2, 3,4,11 months,
Hepatitis B	Birth (risk groups), 3,5,11 months	Birth (risk group), 2,4, 11 months	Birth risk groups) ,2,3,4 months		Birth (at risk), 1-2, 3,4,11 months
Pneumococcal disease	3,5,11 months, PCV13+PPSV23 $\geq$ 65 yrs	2,4, 11 months	PCV13: 2,4,12 months PPSV23 $\geq$ 65 yrs	PCV13: 3,5,12 months, PPSV23 $\geq$ 65 yrs	1-2, 4,11 months
Meningococcal disease	B: 3,4,6,13 months C: 14-15 months MCV4: 12-18 years	C: 4,12 months, 12years	B: 2,4,12months C: 12 months MCV4: 13-15 yrs, 17-25 yrs		MCV4: 14 months, 14 years
Measles	13-15 months, 6 years	12 months, 3-4 years	12 months, 3 years, 10-16 years (catch-up)	15 months, 4 years	14 months, 9 years
Mumps	13-15 months, 6 years	12 months, 3-4 years	12 months, 3 years, 10-16 years (catch-up)	15 months, 4 years	14 months, 9 years
Rubella	13-15 months, 6 years	12 months, 3-4 years	12 months, 3 years, 10-16 years (catchup)	15 months, 4 years	14 months, 9 years
Varicella	13-15 months, 6 years	15 months, 3-4 years, 12 years (catch-up)	Healthy susceptibel household for IC patients (all ages)		
Human papilloma virus infection	12-19 years	12 years	12, 13 years	12 years 13-18 catch-up	12 years

Influenza	Trivalent: ≥ 65 years and risk groups	Trivalent: ≥ 65 years and risk groups	LAIV: 2-7 years Trivalent: risk groups and ≥65	Trivalent: ≥ 65 years and risk groups	> 60 years
Herpes Zoster	50-64 specific groups ≥ 65 years		≥70 years		

Operationalization

Vaccinations will be obtained from the databases by using names of vaccines and database specific codes. Vaccines will be categorized into vaccine types by vaccine preventable disease and type of vaccine (e.g. acellular/whole cell, and type of strains). Brand data will be obtained from the recorded data where available. Variables in the common vaccine input file are: coded patient identifier to link with population, date of administration, vaccine type (antigen code), brand, ATC code (if available), antigen, recorded dose (if available), derived dose (imputed based on chronological order and age/timing of adminstration).

7.7.2 Other Variables and Operationalizations

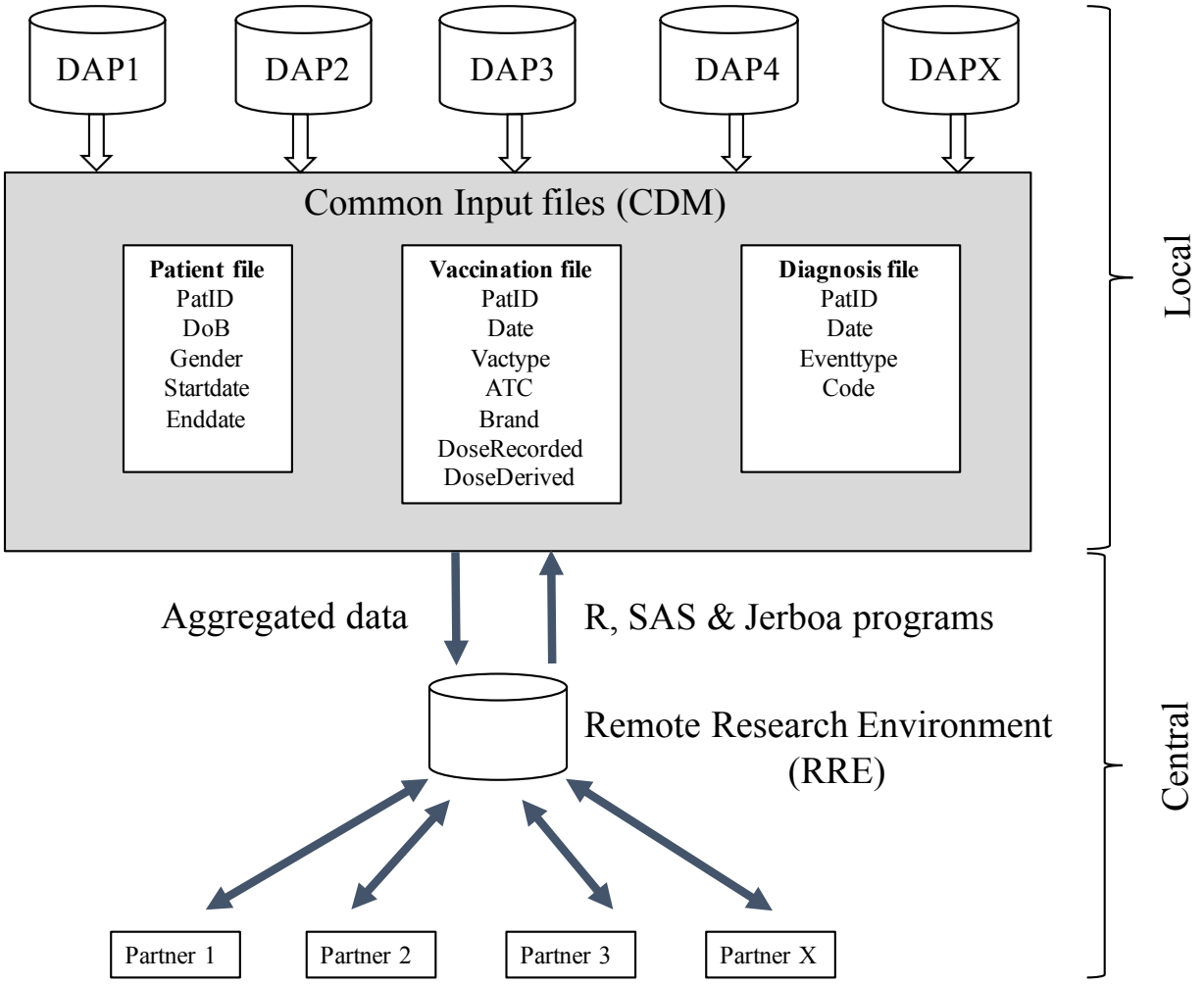
Exposures to types and doses of vaccines will be stratified by database

Dose specific coverage rates will be stratified by birth year by antigen and database and expressed by age as follow-up

7.8 Data Management

7.8.1 Data Processing Steps

Processing of data from the different databases will be using the following steps as per ADVANCE policy (see figure 1):



DAP: data access providers

Figure 1: ADVANCE data management workflow [7]

- Extraction of specific de-identified data from the original databases into specific common input files (population & vaccines). Data access providers are responsible for this task and will conduct this using local software.
- Transformation of the data into analytical datasets suitable for statistical analysis. This will be done locally according to vaccine fingerprint script with one common R script.
- Transfer of aggregated data (coverage estimates) to the remote research environment
- Data pooling and analysis at the remote research environment (RRE).

[7] 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)

- Data pooling across databases will be conducted by the statisticians on the RRE. The RRE has R, SAS, and other programs.

7.8.2 The RRE

The RRE is hosted on an application server (Windows Server 2008R2) located in the data center of the Erasmus University Medical Center (EMC). The data center is a Tier level III data center which means it has multiple independent distribution paths serving the IT equipment and has an expected availability of 99.9%. The server is secured by the EMC firewall and will not have any direct connections to the LAN of the hosting institute. Access to application server is only allowed from a restricted set of IP addresses using two-factor authentication with a password and token. The infrastructure is monitored by the Erasmus MC Computer Emergency Response Team (CERT). Procedures have been developed to ensure data protection and secure file transfer from and to the collaborating partners. The following paragraphs describe these procedures in more detail.

For the RRE many policies for data security have been put in place, for example:
To acquire access to the RRE, each user has to fill in a request form and sign a confidentiality agreement. WP5 leaders (or Steering Committee) need to formally approve each request.

- Users will only have access to the RRE using a remote desktop session.
- Authentication of users consists of two factors: in addition to the basic authentication procedure (with username and password), an authentication with a personal token is performed (SafeNet eToken Pro, www.safenet-inc.com).
- All log on/log off operations are automatically logged (registered) by the RRE.
- The authentication of users is performed by asking, at each login attempt, the username and password (i.e. saved credentials are not allowed).
- Users only gain access to folders/files that are part of the project in which they collaborate. The system administrators can grant permissions to users based on their role in the project.
- Users will not have access to the control panel, internet, and administrative tools.
- Any attempt to copy and paste files between the remote session and local PCs of partners will be disabled.
- All devices on local PCs of partners (i.e. printers, storage...) will be disabled in the remote session.
- A complete log of all requests for files and copies of these files sent outside the RRE will be kept and can be inspected upon request.
- A screensaver will be activated on the remote desktop if the user is not active for a predefined time interval.

Any misconduct or violation of RRE security principles will be notified to the data manager and project manager immediately. Standard operating procedures for access and file transfer rights will be developed within the consortium.

7.8.3 Local data transformation

Data transformation on the basis of the input files will follow the following standard steps that have been defined in prior ADVANCE studies[8,9]

- The population input file of the CDM will be extended with additional variables to define the study population and the time at follow-up. Persons for which person time (calculated as end of follow-up date minus start of follow-up date) is negative or zero will be discarded.
- Cleaning the vaccination CDM input file

    - 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 $Date_{P2} < Date_{P1}$ + 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

7.8.4 Data analysis

Final data analysis and pooling will be conducted inside the RRE, i.e. the user logins in the server and has access to a number of analysis and word processing. All users in the project or WP have access to the data folders (read only) or can share files with others using their personal Share. In the Share, only the owner has write-access; all others only have read permissions. The folder structures can easily be customized to address specific project needs.

The following outcome parameters will be estimated:
Number of persons and persontime by database
Number of doses administered by vaccine and brand during study period

Coverage curves by database by birth year (2000-2017 for childhood vaccines) and birthyears 1930-1960 for vaccines in elderly, per dose. Estimates will be provided at certain ages (see table 5) based on the cumulative distribution functions

[8] 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)

[9] 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)

For the calculation of coverage we will follow the following calculations

We assign a letter (A, B, C, D, E) to every age week of every person.

$$A_i = in\ follow-up\ (FU)\ during\ age\ i, vaccinated\ during\ age\ i$$
$$B_i = in\ FU\ during\ age\ i, vaccination\ recorded\ before\ age\ i$$
$$C_i = in\ FU\ during\ age\ i, no\ recorded\ vaccination\ before\ age\ i$$
$$D_i = Not\ in\ FU\ during\ age\ i, vaccination\ recorded\ before\ age\ i$$
$$E_i = Not\ in\ FU\ during\ age\ i, no\ recorded\ vaccination\ before\ age$$

From the data aggregated by birthyear, we produce the following estimators;

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)

We estimate the cumulative probability density (Φ_A) 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. We interpreted the increase between week $i-1$ ($= \Phi_A(t_{i-1})$) and i ($= \Phi_A(t_i)$) 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. We use a 5000-step numerical integration to quantify MFU for each age-week. $\Phi_A(t_i)$ 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.

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

We subsequently multiply the meaningful follow-up for week i 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 we need to normalize the proportion of meaningful follow-up at the end of week i;

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

Finally we weight the total number of vaccinations at the end of week i by the normalized $MFU_{proportion,i}$.

$$CDF_i = \frac{\sum_{0 \to i} A_i}{MFU_{proportion.normalized,i}}$$

The age at which the coverage will be estimated is provided in table 5

Table 5: Age at which we will estimate coverage for reporting and comparisons using CDF

Vaccine	D1	D2	D3	Booster 1	Booster 2	Elderly	WHO assessment
Measles	24 months	10 years					MCV1
Mumps	24 months	10 years					MCV1
Rubella	24 months	10 years					MCV1
Diphteria	12 months	12 months	24 months	8 years	18 years		DTP1, 3
Tetanus	12 months	12 months	24 months	8 years	18 years		DTP1, 3
Pertussis	12 months	12 months	24 months	8 years	18 years		DTP1, 3
HIb	12 months	12 months	24 months	8 years			Hib3
Hepatitis B	12 months	12 months	24 months				HepB3
Polio	12 months	12 months	24 months	8 years	18 years		Pol3
PCV	12 months	12 months	24 months			Age categories	PCV3
Influenza	8 years					Yearly, >65	
Varicella	24 months	8 years					
HPV	18 years	18 years					
Herpes Zoster						Age categories	
Rotavirus	12 months	12 months	12 months				RotaC (2nd or third)

Coverage: charts with weekly 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 birthcohort j will be calculated from each database. This will be done for each database, type of vaccine.

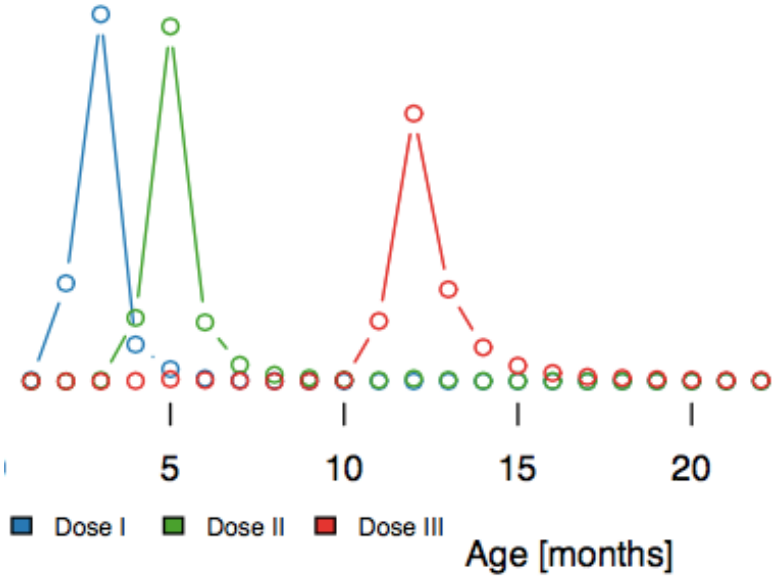


Figure 2: Example of histogram plot with the number of doses by age

Coverage: line plots with vaccination coverage (%) by year-month birth cohorts over time
For every birth-year cohort, we will evaluate the vaccination coverage by dose over age for each estimation method. The coverage at week i for birth 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. In addition we will provide coverage estimate by cumulative incidence, and cumulative distribution function

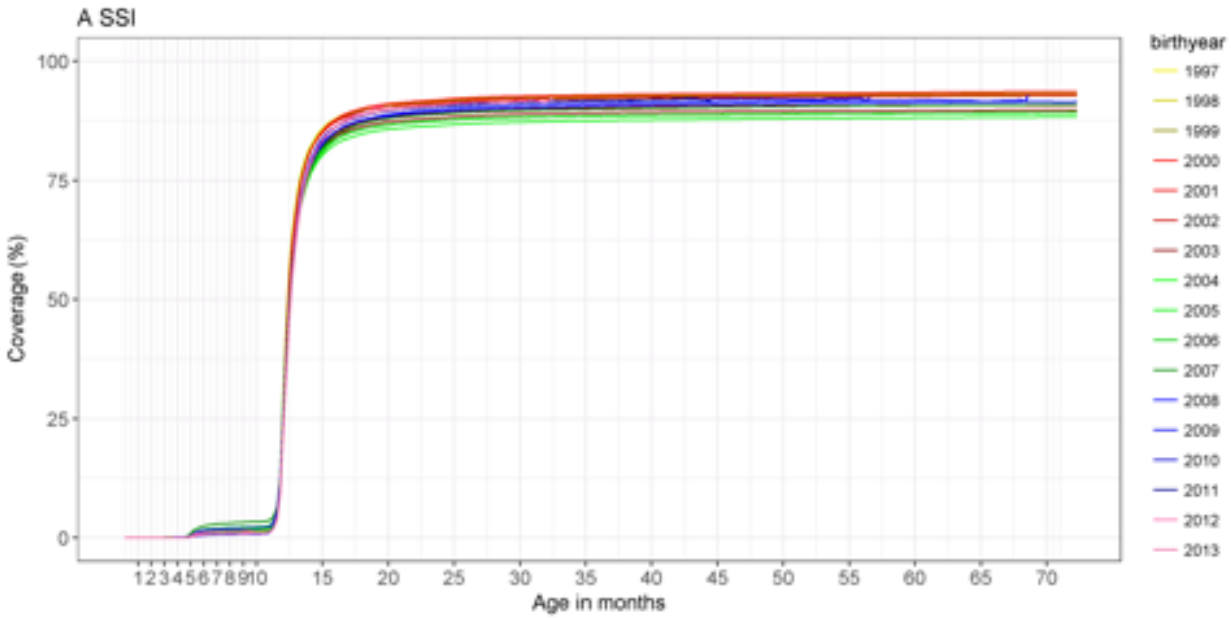


Figure 3: Example of coverage curves

Benchmarking

In order to assess completeness of vacciation data in the specific databases, we will compare the coverage estimates obtained from the cumulative distribution function at specific ages for the most recent calendar year and birthcohort with enough follow-up with published estimates from WHO and/or literature or national public health datasources.

No specific testing will be done. A priori: if coverage estimates in the databases deviate more than 10% (relative) from published/WHO data, vaccination data will be considered incomplete. It is of note that coverage reported to WHO may have varying origin, birthcohorts, years of assessment, described in the links below. We will use the most recent data available.

Spain reference data WHO

https://www.who.int/immunization/monitoring_surveillance/data/esp.pdf

United Kingdom reference data WHO

https://www.who.int/immunization/monitoring_surveillance/data/gbr.pdf

Dutch reference data WHO

https://www.who.int/immunization/monitoring_surveillance/data/nld.pdf

Danish reference data WHO

https://www.who.int/immunization/monitoring_surveillance/data/dnk.pdf

Italian reference data WHO

https://www.who.int/immunization/monitoring_surveillance/data/ita.pdf

For HPV vaccination coverage we will use the below reference data

https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(16)30099-7/fulltext

For Influenza we will use data from the VENICE consortium as far as possible and benchmark data only for elderly as risk groups will not be identified as part of this protocol.

https://ecdc.europa.eu/sites/portal/files/documents/influenza-vaccination-2007–2008-to-2014–2015.pdf

8. Quality Control

8.1 Handling of missing data

Incompleteness of vaccination data will be assessed using national benchmarks on coverage.

8.2 Quality check and benchmarking

Locally data will be extracted by the data controllers using local programs. An iterative process is used to verify extractions between the study team and data access provider, when the aggregate data are transferred to the RRE

All data transformation programs (estimation exposure and coverage) will be centrally programmed per agreed coding standards. R version 3.4.0 will be used for transformation and statistical analyses. The POC 2 programming file to transform the CDM files to the analytical datasets is a modification of existing programs to transform the CDM files to the analytical datasets used for the coverage study. The programs were double coded in R and SAS. For the POC-2 adaptations to the programs will be quality controlled by a second statistician. Benchmarking of coverage rates will be done against the national/regional statistics or data from the VENICE consortium or sales data (in absence of other suitable data)

8.3 Record Retention

Documents that individually and collectively permit evaluation of the conduct of a project and the quality of the data produced will be retained for a period of 5 years in accordance with Good Participatory Practice (GPP) guidelines. These documents should be retained for a longer period, however, if required by the applicable regulatory requirements or by an agreement between project partners. It is the responsibility of the principal investigator to inform the other investigators/institutions as to when these documents no longer need to be retained.

Project records or documents may also include the analysis files, syntaxes (usually stored at the site of the database), and questionnaires.

8.4 Advisory Committee

None

8.5 Use of the data generated in this project

The aggregated data & graphics generated in this project may be used for publication and for interactive display on the VAC4EU website, through controlled access. All data access providers will be asked to participate in publications and will be asked to sign a datasharing agreement and conditions of use/access for interactive display.

9. Protection of human subjects

9.1 Regulatory and Ethical Compliance

The European legislation describes obligations to be fulfilled by marketing authorisation holders (MAHs) and national competent authorities for medicinal products (including vaccines) authorised in the European Union (EU). The European legislation does not apply to post-authorisation studies conducted by organisations such as academia, medical research charities or research organisations in the public sector. These organisations should follow local requirements defined in the national legislation applicable in the countries where research is conducted. In the context of the ADVANCE consortium, proof-of-concept studies (POCs) will be conducted to test new approaches (data sources, methods) by using a test case; within this framework the POC studies will collect and use data on vaccines authorised in the EU. Vaccine MAHs are partners of the ADVANCE consortium and participate in the design, conduct and funding (through in-kind contribution) of the POCs concerning vaccines (at the product or substance level) for which they hold an authorisation. They should therefore be considered as having a control on the design of POCs, in which case requirements of the GVP applies. (21)
The GVP requirements' will be addressed in the following ways:
1) The project proposal and project report will be posted on the EU PAS register.
2) The POC-2 will be monitored by MAHs as PAS.
3) This project aims to test availability of vaccine data to create readiness for benefit-risk monitoring of vaccines in Europe, more specifically:
The objectives of this project are on methodological aspects and feasibility and not intended to provide any information on the safety of the vaccines. Therefore this project is not considered as a PASS.
4) Management and reporting of adverse events/adverse reactions:
This project is observational, based on secondary use of data in large healthcare databases and will not look at events following vaccination. Thus the reporting of suspected adverse reactions in the form of individual case safety reports (ICSRs) is not required and no individual adverse events/reactions will be summarised in the final project report.

9.2 Informed Consent

Databases with an internal review board approval indicating that informed consent is waived and the rational for this decision will be included in the analyses. Informed consent is obtained for the Pedianet database to allow linkage with an outside register.

9.3 Responsibilities of the Investigator and IRB/IEC/REB

The proposal and waiver of informed consent and/or ethical clearance must be reviewed and approved by a properly constituted institutional review board/independent ethics committee/research ethics board (IRB/IEC/REB) before project start. A signed and dated statement that the proposal has been approved by the IRB/IEC/REB and waiver of informed consent must be given by the data access providers to the principal investigator before project initiation.

9.4 Proposal Adherence

Investigators will apply due diligence to avoid proposal deviations. If the investigator feels a change to the proposal would improve the conduct of the project this must be considered a proposal amendment, and unless such an amendment is agreed upon by all partners involved and approved by the IRB/IEC/REB it cannot be implemented. All significant proposal deviations will be recorded and reported in the Project Report.

10. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ ADVERSE REACTIONS

N/A

11. PLANS FOR DISSEMINATING AND COMMUNICATING RESULTS

11.1 Registration in Public Database(s)

Coordinators assure that the key design elements of this proposal will be posted in the EU PAS register in compliance with current regulations.
Coordinators also assure that key results of this proposal will be posted in a publicly accessible database within the required time-frame from completion of the data collection where applicable and in compliance with current regulations.

11.2 Publications

Further to legislated data disclosure, the results of this project will be published as scientific papers in peer-reviewed journals. Preparation of such manuscripts will be prepared independently by the investigators and in accordance with the current guidelines of STrengthening the Reporting of OBservational studies in Epidemiology (STROBE). The ADVANCE Steering Committee will be entitled to view the results and interpretations included in the manuscript and provide comments prior to submission of the manuscript for publication.

12. Timelines

August 2018	Outline to SC
September	Discussion in GAM
December	Protocol submission for in house clearance
December –January	Conversion of data into common data model locally (feasibility)
February 2019	Data transformation & sharing of results
31, March 2019	Report delivered

Annex 1: EncePP Checklist

Doc.Ref. EMA/540136/2009

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: Exposure and coverage to routine schedule vaccines in different EU countries

EU PAS Register® number: 27851
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[10]	☒	☐	☐	Page 8
1.1.2 End of data collection[11]	☒	☐	☐	Page 8
1.1.3 Progress report(s)	☒	☐	☐	Page 8
1.1.4 Interim report(s)	☒	☐	☐	Page

[10] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[11] Date from which the analytical dataset is completely available.

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	Page 26
1.1.6 Final report of study results.	☒	☐	☐	Page 8

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	Ch. 5
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	Ch. 5
2.1.2 The objective(s) of the study?	☒	☐	☐	Ch. 6
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	Ch7.4
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☒	☐	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☒	☐	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	Ch 7.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	Ch 7.3
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	Ch 7.8.4
3.4 Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☐	☐	☒	
3.5 Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☒	☐	☐	Ch 10

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	Ch 7.4
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	Ch 7.5/7.6
	4.2.2 Age and sex	☒	☐	☐	Ch 7.5
	4.2.3 Country of origin	☒	☐	☐	Ch 7.3
	4.2.4 Disease/indication	☐	☐	☒	
	4.2.5 Duration of follow-up	☒	☐	☐	Ch 7.6
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	Ch 7.5

Comments:

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
5.1	Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	7.7
5.2	Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	7.8.4
5.3	Is exposure categorised according to time windows?	☐	☒	☐	
5.4	Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	7.8.4
5.5	Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6	Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	7.8.4

Comments:

Section 6: Outcome definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
6.1	Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	7.8.4

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	7.8.4
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☐	☐	☒	
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☒	☐	

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☐	☐	☒	
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☐	☐	☒	
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	7.8.4

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☐	☒	

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	7.3
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☐	☐	☒	

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1.3 Covariates and other characteristics?	☒	☐	☐	7.3
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	7.3
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	7.3
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	7.3
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	7.3
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	7.3
9.3.3 Covariates and other characteristics?	☒	☐	☐	7.3
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	7.3

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	7.8.4
10.2 Is study size and/or statistical precision estimated?	☐	☐	☒	
10.3 Are descriptive analyses included?	☒	☐	☐	7.8.4
10.4 Are stratified analyses included?	☒	☐	☐	7.8.4
10.5 Does the plan describe methods for analytic control of confounding?	☐	☐	☒	
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☐	☒	
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	8.1
10.8 Are relevant sensitivity analyses described?	☐	☐	☒	

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	7.8
11.2 Are methods of quality assurance described?	☒	☐	☐	8.2
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☐	☐	☒	
12.1.2 Information bias?	☒	☐	☐	7.8.4
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☐	☐	☒	
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☐	☒	☐	

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	9.1
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	7.8

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	4

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	11
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	11

Comments:

Name of the main author of the protocol: Miriam Sturkenboom

Date: 01/02/2019

Signature: