VAC4EU

VAccine monitoring Collaboration for Europe

Proposal

Title: Rapid assessment and self-controlled risk intervals studies of Moderna COVID-19 vaccine

Table of Contents

1. Background

1.1 Phase III clinical trial on Moderna vaccine

The Moderna COVID-19 Phase 3 vaccine efficacy trial, known as COVE, was begun under Operation Warp Speed (OWS), a multi-agency collaboration led by HHS and the Department of Defence that aims to accelerate the development, manufacturing and distribution of medical countermeasures for COVID-19. The study is A Phase 3, Randomized, Stratified, Observer-Blind, Placebo-Controlled Study to Evaluate the Efficacy, Safety, and Immunogenicity of mRNA-1273 SARS-CoV-2 Vaccine in Adults Aged 18 Years and Older. The phase III trial is registered at clinicaltrials.gov and search identifier NCT04470427

The mRNA-1273 vaccine candidate was co-developed by the Cambridge, Massachusetts-based biotechnology company Moderna, Inc., and the National Institute of Allergy and Infectious Diseases (NIAID), part of the National Institutes of Health. It combines Moderna's mRNA (messenger RNA) delivery platform with the stabilized SARS-CoV-2 spike immunogen (S-2P) developed by NIAID scientists.

More than 30,000 participants at 100 clinical research sites in the United States are participating in the phase III study, which launched on July 27, 2020, after results from earlier stage clinical testing indicated that the vaccine candidate is well-tolerated and immunogenic.

Figure 1: Overview of efficacy data studies (from lshtm vaccine tracker)

COVE study Inclusion Criteria[1]:

- Participants who are at high risk of SARS-CoV-2 infection, defined as adults whose locations or circumstances put them at appreciable risk of exposure to SARS-CoV-2 and COVID-19.
- Understands and agrees to comply with the study procedures and provides written informed consent.
- Able to comply with study procedures based on the assessment of the Investigator.

[1] https://clinicaltrials.gov/ct2/show/study/NCT04470427

- Female participants of non-childbearing potential may be enrolled in the study. Non-childbearing potential is defined as surgically sterile (history of bilateral tubal ligation, bilateral oophorectomy, hysterectomy) or postmenopausal (defined as amenorrhea for ≥12 consecutive months prior to Screening without an alternative medical cause). A follicle-stimulating hormone (FSH) level may be measured at the discretion of the Investigator to confirm postmenopausal status.
- Female participants of childbearing potential may be enrolled in the study if the participant fulfils all the following criteria:
    - Has a negative pregnancy test at Screening and on the day of the first dose (Day 1).
    - Has practiced adequate contraception or has abstained from all activities that could result in pregnancy for at least 28 days prior to the first dose (Day 1).
    - Has agreed to continue adequate contraception through 3 months following the second dose on Day 29.
    - Is not currently breastfeeding.
- Male participants engaging in activity that could result in pregnancy of sexual partners must agree to practice adequate contraception and refrain from sperm donation from the time of the first dose and through 3 months after the second dose.
- Healthy adults or adults with pre-existing medical conditions who are in stable condition. A stable medical condition is defined as disease not requiring significant change in therapy or hospitalization for worsening disease during the 3 months before enrolment.

Exclusion Criteria:

- Is acutely ill or febrile 72 hours prior to or at Screening. Fever is defined as a body temperature ≥38.0°C/100.4°F. Participants meeting this criterion may be rescheduled within the relevant window periods. Afebrile participants with minor illnesses can be enrolled at the discretion of the Investigator.
- Is pregnant or breastfeeding.
- Known history of SARS-CoV-2 infection.
- Prior administration of an investigational coronavirus (SARS-CoV, Middle East Respiratory Syndrome [MERS]-CoV) vaccine or current/planned simultaneous participation in another interventional study to prevent or treat COVID-19.
- Demonstrated inability to comply with the study procedures.
- An immediate family member or household member of this study's personnel.
- History of anaphylaxis, urticaria, or other significant adverse reaction requiring medical intervention after receipt of a vaccine.
- Bleeding disorder considered a contraindication to intramuscular injection or phlebotomy.
- Has received or plans to receive a vaccine within 28 days prior to the first dose (Day 1) or plans to receive a non-study vaccine within 28 days prior to or after any dose of investigational product (except for seasonal influenza vaccine).
- Has participated in an interventional clinical study within 28 days prior to the day of enrolment.
- Immunosuppressive or immunodeficient state, including human immunodeficiency virus (HIV) infection, asplenia, and recurrent severe infections.
- Has received systemic immunosuppressants or immune-modifying drugs for >14 days in total within 6 months prior to Screening (for corticosteroids ≥20 milligram (mg)/day of prednisone equivalent). Topical tacrolimus is allowed if not used within 14 days prior to Screening.
- Has received systemic immunoglobulins or blood products within 3 months prior to the day of Screening.
- Has donated ≥450 millilitres (mL) of blood products within 28 days prior to Screening.

The COVE trial is sponsored by Moderna. Twenty-five NIH-funded COVID-19 Prevention Network sites enrolled participants in the trial.

Primary outcomes are

- Number of Participants with a First Occurrence of COVID-19 Starting 14 Days after Second Dose of mRNA-1273 [Time Frame: Day 29 (second dose) up to Day 759 (2 years after second dose)]
- Number of Participants with Adverse Events (AEs) or Medically Attended AEs (MAAEs) Leading to Withdrawal [Time Frame: Up to Day 759 (2 years after second dose)]
- Number of Participants with Solicited Local and Systemic Adverse Reactions (ARs) [Time Frame: Up to Day 8 (7 days after first dose) and up to Day 36 (7 days after second dose)]
- Number of Participants with Unsolicited AEs [Time Frame: Up to Day 57 (28 days after each dose)]

Secondary outcomes are:

- Number of Participants with a First Occurrence of Severe COVID-19 Starting 14 Days after Second Dose of mRNA-1273 [Time Frame: Day 29 (second dose) up to Day 759 (2 years after second dose)]
- Clinical signs indicative of severe COVID-19 as predefined for the study.
- Number of Participants with a First Occurrence of Either COVID-19 or SARS-CoV-2 Infection regardless of symptomatology or Severity Starting 14 Days after Second Dose of mRNA-1273 or Placebo [Time Frame: Day 29 (second dose) up to Day 759 (2 years after second dose)]]
- Clinical signs indicative of COVID-19 and SARS-CoV-2 Infection as predefined for the study.
- Number of Participants with a Secondary Case Definition of COVID-19 Starting 14 days after Second Dose of mRNA-1273 or Placebo [Time Frame: Day 29 (second dose) up to Day 759 (2 years after second dose)]
- Clinical signs indicative of secondary case definition of COVID-19 as predefined for the study.
- Number of Participants with a First Occurrence of COVID-19 Starting 14 days after First Dose of mRNA-1273 or Placebo [Time Frame: Day 1 (first dose) up to Day 759 (2 years after second dose)]
- Clinical signs indicative of COVID-19 as predefined for the study.
- Number of Participants with a First Occurrence of COVID-19 Starting 14 days after Second Dose of mRNA-1273 or Placebo regardless of evidence of prior SARS-CoV-2 Infection [Time Frame: Day 29 (second dose) up to Day 759 (2 years after second dose)]
- Clinical signs indicative of COVID-19 and SARS-CoV-2 infection as predefined for the study.
- Number of Participants with a First Occurrence of SARS-CoV-2 Infection in the Absence of Symptoms Defining COVID-19 Starting 14 days after Second Dose of mRNA-1273 or Placebo [Time Frame: Day 29 (second dose) up to Day 759 (2 years after second dose)]
- Clinical signs indicative of COVID-19 and SARS-CoV-2 infection as predefined for the study.
- Geometric Mean Titre (GMT) of SARS-CoV-2 Specific Neutralizing Antibody (nAb) [Time Frame: Day 1, Day 29, Day 57, Day 209, Day 394, and Day 759]
- Geometric Mean Fold Rise (GMFR) of SARS-CoV-2 Specific nAb [Time Frame: Day 1, Day 29, Day 57, Day 209, Day 394, and Day 759]
- Quantified Levels or GMT of S Protein-Specific Binding Antibody (bAb) [Time Frame: Day 1, Day 29, Day 57, Day 209, Day 394, and Day 759]
- GMFR of S Protein Specific bAb [Time Frame: Day 1, Day 29, Day 57, Day 209, Day 394, and Day 759]

Missing data

- No data on Europeans
- Lack of data on vulnerable populations: children, pregnant women and immune compromised subjects as well as people with prior COVID-19 disease and those with history of anaphylaxis
- Impact of longer duration of use, currently follow-up is about 2-3 months with lacking information on longer term use
- Lack of power to detect AESI.
    - A total sample size of 15,000 treated vaccine recipients would be able to rule out at 95% confidence events with a frequency of 1/5000
    - AESI are typically serious and rare with most frequencies ranging between 1/10,000-1/100,000, which typically cannot be detected in the RCTs

1.2 Europe and COVID-19 vaccines

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

- 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

It is unknown which potential vaccine, if any, will successfully complete the development and authorisation process and thus meet efficacy and safety criteria to be placed on the EU market, but based on recent interim analyses of pivotal phase III trials that show high efficacy, Pfizer was given emergency use rights in in the UK on December 2, and a license from EMA on December 21, 2020. Moderna vaccine is scheduled for December 29 or January 2021.

The WHO-SAGE committee developed a values framework for the allocation and prioritization of COVID-19 vaccination. 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 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

[2] https://apps.who.int/iris/bitstream/handle/10665/334299/WHO-2019-nCoV-SAGE_Framework-Allocation_and_prioritization-2020.1-eng.pdf?ua=1&ua=1

ongoing clinical trials, the following groups seem likely to be addressed first, but this will be decided in each country separately in the coming weeks:

- 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)
- Sociodemographic groups at disproportionately higher risk of severe disease or death

A recognized challenge of the Pfizer vaccine is the cold chain requirements, which may not be feasible in all settings. This vaccine is currently targeted in most countries for health care workers, whereas Moderna vaccine that may stored in the fridge might be targeted at at-risk populations that are vaccinated in primary care.

1.3 Adverse events of special interest (AESI)

The ACCESS project (vACcine Covid-19 monitoring readinESS) which is funded by EMA to prepare Europe for monitoring COVID-19 vaccines has established a list of 38 conditions as adverse events of special interest (AESI), which has been endorsed by EMA (table 1).

Table 1 AESI as defined in the ACCESS project

Body system / Classification	AESI
Auto-immune diseases	Guillain-Barré Syndrome
	Acute disseminated encephalomyelitis
	Narcolepsy
	Acute aseptic arthritis
	Diabetes (type 1 and broader)
	(Idiopathic) Thrombocytopenia
Cardiovascular system	Acute cardiovascular injury including: Microangiopathy, Heart failure, Stress cardiomyopathy, Coronary artery disease, Arrhythmia, Myocarditis
Circulatory system	Coagulation disorders: Thromboembolism, Hemorrhage
	Single Organ Cutaneous Vasculitis
Hepato-gastrointestinal and renal system	Acute liver injury
	Acute kidney injury
Nerves and central nervous system	Generalized convulsion
	Meningoencephalitis
	Transverse myelitis
Respiratory system	Acute respiratory distress syndrome

Skin and mucous membrane, bone and joints system	Erythema multiforme
	Chilblain – like lesions
Other system	Anosmia, ageusia
	Anaphylaxis
	Multisystem inflammatory syndrome in children
	Death (any causes)
	COVID-19 disease (by levels of severity): Level 1: any recorded diagnosis, level 2: hospitalization for COVID-19 (confirmed or suspected), level 3: ICU admission in those with COVID-19 related admission; level 4: Acute respiratory distress requiring ventilation (ARDS) during a hospitalization for COVID-19; level 5 death during a hospitalization for COVID-19 (any cause)
	Sudden death
Pregnancy outcome - Maternal	Gestational Diabetes
	Preeclampsia
	Maternal death
Pregnancy outcome - Neonates	Fetal growth restriction
	Spontaneous abortions
	Stillbirth
	Preterm birth
	Major congenital anomalies
	Microcephaly
	Neonatal death
	Termination of Pregnancy for Foetal Anomaly

2. VAC4EU

2.1 Overview

The Vaccine monitoring Collaboration for Europe (VAC4EU) https://vac4eu.org is a non-for-profit international association registered in Belgium. It was initiated as a sustainable solution of the Innovative Medicines Initiative funded ADVANCE project, which had 47 partners across academics, vaccine manufacturers, public health organizations and regulatory agencies, including the European Medicines Agency (EMA) and the European Centre for Disease prevention and Control. VAC4EU implements the blueprint for a vaccine monitoring system that was written by ECDC for the ADVANCE project. VAC4EU vision is best scientific evidence on vaccine coverage, benefits and risks in Europe to support data-driven decision making. The mission of VAC4EU is to access, characterize and analyse available and newly collected health data to allow for evidence-based decisions by people, who need to either regulate, advise, prescribe or decide on vaccines (https://vac4eu.org). Its statutes

are publicly available through the following link
Since its registration in January 2020, 20 organizations (research and public health organizations) across Europe have joined VAC4EU, subscribing to the vision and mission and recognizing the need to work together to monitor vaccines. Any of these organizations can act as coordination center, PI and/or data access/expertise provider.

2.2. Members

V
AC4EU is a network organization registered in the ENCePP database
http://www.encepp.eu/encepp/viewResource.htm?id=37277

Members are recognized centers of excellence by the European Network of Excellence in Pharmacoepidemiology and Pharmacovigilance according to EMA. Participating members (as of December 20, 2020) are:

1. **University Medical Centre Utrecht**, the Netherlands, academic medical center with ample expertise in vaccine studies, pre-licensure and post-licensure. Host of the VAC4EU secretariat. Coordinator of the ACCESS project. https://www.umcutrecht.nl/nl ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=38449
2. **Sciensano**, national public health agency Belgium, with ample expertise in vaccine studies, pre-licensure and post-licensure. https://www.sciensano.be
3. **LAREB,** Netherlands national pharmacovigilance center, https://www.lareb.nl. Ample expertise in regular pharmacovigilance and intensive monitoring of influenza and pneumococcal vaccines www.lareb.nl, ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=14143
4. **RIVM**, the Dutch national public health institute, ample expertise in vaccine coverage, effectiveness and safety studies with primary data collection and secondary use of data. https://www.rivm.nl
5. **Leibniz Institute for Prevention Research and Epidemiology** – BIPS, Germany. Home to the German Pharmacoepidemiologic Research Database (GePaRD), claims data from statutory health insurance providers and currently includes information on about 25 million persons. https://www.bips-institut.de , ENCePP http://www.encepp.eu/encepp/viewResource.htm?id=26905
6. **Societa Servizi Telematici,** Italy. Home to the PEDIANET database covering primary care pediatricians in Italy http://pedianet.it/en/primary-health-care-project . ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=20131
7. **Bordeaux PharmacoEpi (BPE)**, France, is a research platform of the Université de Bordeaux specialized in real world evidence with extensive experience in conducting fields studies and studies based on the French national healthcare database (SNDS). https://www.bordeauxpharmacoepi.eu and in ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=38744
8. **Research Triangle Institute Health Solutions (RTI-HS), Spain** a private non-for-profit research organization. Pharmacoepidemiology and Risk Management expertise including EMA-mandated safety and utilization studies. https://www.rtihs.org/ . ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=36190
9. **Lazio Regional Health Service,** Italy. Regional public health agency. Expertise in vaccine studies and COVID-19 studies, access to administrative health care data (mortality, hospital admissions, ER access, drug claims & vaccines, health care assistance), in Lazio region (6 million residents). https://www.deplazio.net/en, http://www.encepp.eu/encepp/viewResource.htm?id=24593
10. **Agenzia Regionale di Sanità (ARS) Tuscany**, Italy. Regional public health agency. Expertise in vaccine studies and COVID-19 studies, access to administrative health care data (mortality, hospital admissions, ER access, drug claims & vaccines, health care assistance), in Tuscany region (6 million residents). https://www.ars.toscana.it, ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=24413
11. **Penta foundation**, a private non-for-profit foundation hosting a global pediatric research network, which has coordinated hundreds of multisite clinical trials and cohort studies in infectious diseases in the past 30 years. https://penta-id.org/who-we-are/compliance/
12. **PHARMO Institute**, the Netherlands. A private research organization conducting studies on use

and effects of medicinal products. Home to the PHARMO data network, is a population-based network from different primary and secondary healthcare settings in the Netherlands. www.pharmo.nl. ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=17375

13. **IDIAP-Jordi Gol**, Spain. A public primary health care research organization. Home to the SIDIAP database, Information System for Research in Primary Care, electronic health data. https://www.idiapjgol.org/index.php/en/. Encepp: http://www.encepp.eu/encepp/viewResource.htm?id=32245
14. **The Foundation for the Promotion of Health and Biomedical Research of Valencia Region**, Spain FISABIO, FISABIO has access to regional health and vaccine data
15. **University Verona**, Italy. An academic research organization and important regional pharmacovigilance center for Italy. ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=21184
16. **University Oslo**, Norway. Public academic research organization. Access to Norwegian health and vaccine data. http://www.encepp.eu/encepp/viewResource.htm?id=15887
17. **Drug Safety Research Unit**, DSRU, UK. The DSRU is a private research organization in the UK which monitors the safety of medicines and vaccines and uses UK databases. ENCePP: http://www.encepp.eu/encepp/viewResource.htm?id=38664
18. **University of Lyon:** academic research organization able to provide access to public and private French databases. ENCePP http://www.encepp.eu/encepp/viewResource.htm?id=15425
19. **University Aarhus:** Academic research organization, having access to national Danish health registers. ENCePP http://www.encepp.eu/encepp/viewResource.htm?id=36219
20. **Royal College for Primary Care physicians (SIMG) Italy:** data access provider to Italian GP network and the Health Search Data network, a GP based data collection. ENCePP http://www.encepp.eu/encepp/viewResource.htm?id=6872

2.3 VAC4EU tools for distributed study conduct

The VAC4EU research infrastructure provides a toolbox and IT services to rapidly conduct studies in a distributed manner where data remain local and the data access provider transforms data in a common data model, that is then analyzed with standard R-tools (Fig. 2). This process was tested and very successful in the IMI-ADVANCE proof near real time monitoring studies. For the currently proposed study we will use two IT dedicated tools to engage the data access providers (DAP) who organize local access to data. The study team will comprise representatives from the DAPs The toolbox to be utilized are available protocol templates from ACCESS, event definitions, codes and algorithms, as well as the R-scripts that were used to assess quality of data and the incidence rates

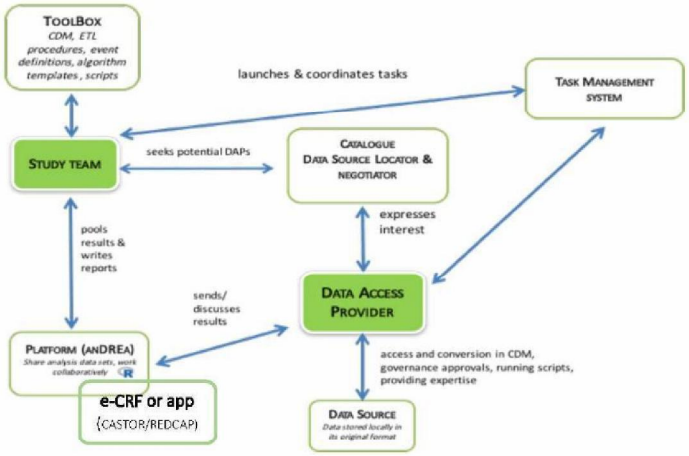

Figure 2: Graphical display how studies can be rapidly conducted using VAC4EU tools

Task management system
This system supports launching and monitoring up of tasks for distributed tasks (e.g. protocol approvals, scripts release)

Remote research environment anDREa platform
The azure Digital Research Environment (anDREa, https://www.andrea-consortium.org/azure-dre/) is a cloud based, globally available research environment where data is stored and organized securely and where researchers can quickly generate workspaces to collaborate in. Within these workspaces, researchers have preinstalled applications at their disposal, as well as the ability to bring own tooling. The DRE facilitates users to collaborate on research projects in a safe, yet flexible computing and storage environment. The architecture of the DRE allows researchers to use a solution within the boundaries of data management rules and regulations. Within the DRE platform each of the projects you are a member of consists of a separate, secure folder, called a 'workspace'. Each workspace is completely secure, so researchers are in full control of their data. Each workspace has its own list of users, which can be managed by its administrators. Each workspace is fully scalable with regard to data quantity and computing power, thereby supporting anything from small to complex multi-center, multisource studies.

2.4 Access to data/expertise in VAC4EU members

The work for this proposal will be executed by a team of highly experienced and dedicated public health, pharmacoepidemiology, data science, statistical and vaccinology experts. Together they join efforts to create the EU infrastructure to monitor COVID-19 vaccines based on their long-lasting and globally recognized experience in pharmacoepidemiology, vaccine coverage, effectiveness and safety research. We describe the expertise of the key scientists and some recent publications.

3. Research question and objectives

The overall aim of the proposed study would be to explore and monitor the safety of **Moderna COVID19** vaccine in a real world setting after Moderna vaccine license in Europe

3.1 Ecological Methods

Primary objective: To assess whether there is increased incidence of AESI following introduction of Moderna vaccine as compared to the period prior to introduction of Moderna vaccine

Secondary objective: To assess whether there is increased incidence of AESI following introduction of Moderna vaccine as compared to the period prior to introduction of a Moderna vaccine in groups defined by age, and at-risk populations

3.2 Unadjusted Self-Controlled Risk Interval Analysis

Primary objectives: To determine whether there is increased incidence of AESI following vaccination with specific Moderna as compared to a pre or post-vaccination control period.

4. Methods

A study protocol, following the format described in GVP Module VIII3, section VIII.B.3.1., would be

provided if Moderna would like to proceed. The protocol development will be led by organizations of choice. Available organizations for protocol development are UMCU and RTI-HS, who also prepared the safety protocols in the ACCESS project for EMA

4.1 Designs proposed

4.1.1 Ecological design

Retrospective, multi-database cohort study to assess changes in the incidence rate of AESI and to evaluate the impact of COVID-19 vaccine introduction on the occurrence of AESI

Ecological analyses employ a simple before/after comparison of incidence rates or make use of multiple time points before and after an intervention in an interrupted time series (ITS) analysis to compare a pre-vaccine introduction period versus a post-vaccine introduction period. In this design there is no need to know vaccination status at an individual level, rather the country/regional introduction date is used. In an ITS analysis, which is one ecological design, slope and/or level in incidence trends over time can be compared in pre- vs. post-intervention time in a regression model (see **Figure** 3.

In an ITS analysis, incidence rates are calculated at regularly spaced intervals (month, quarter) and the study period should include a sufficient number of time intervals prior to the vaccine introduction to allow for assessment of temporal trends, independent of vaccination. Autocorrelation should be assessed and controlled for if present (Bernal, 2017). Power in ITS analyses is dependent not only upon the number of time points, but upon the sample size at each time point, so these should be taken into balanced consideration.

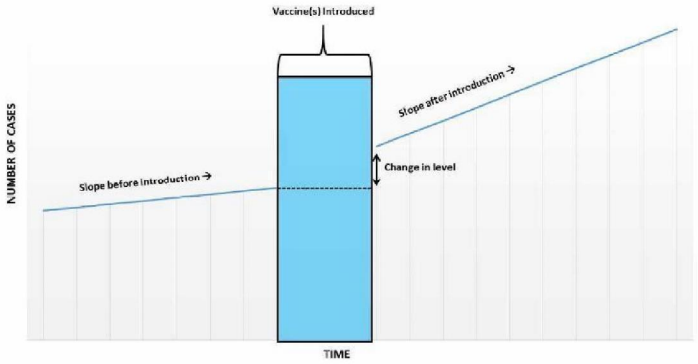

Figure 3. Interrupted Time Series Design adapted from Ramsey, 2013

4.1.2 Self-Controlled Risk Interval (SCRI)

Retrospective (multi)-database case-only study including subjects who were vaccinated with Moderna vaccine and experienced AESI (figure 4)

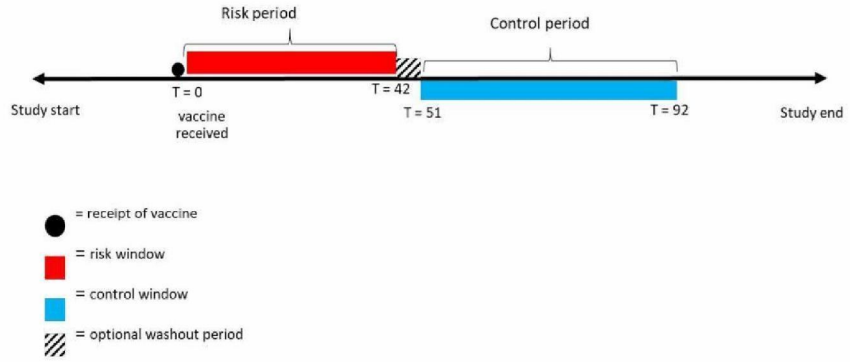

Figure 4. Self-Controlled Risk Interval design

Note: Example with a risk period of 42 days and a control period of 42 days

If needed a comparative cohort study can also be performed.

4.2 Setting & feasibility

For the implementation of this study we would use electronic health care databases. This will require that data sources are/can

- Population based
- periodically update their medical data (diagnoses, with short lag times for key data (hospitalizations, vaccinations, population, medicines)
- are willing/able to participate with Moderna as sponsor
- provide background rates as in the ACCESS BGR protocol (EU PAS before participation, to benchmark and have proper background rates using a distributed approach and common data model
- Are able to identify vaccine brand

VAC4EU data sources are enlisted in Appendix 1.

4.3 Study population & study period

4.3.1 Source Population

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

4.3.2 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 (such as study period, design requirements and exclusion of prior events.

Eligible individuals will be identified in each of the database using a pre-specified selection process

and/or by applying pre-specified algorithms and attrition diagrams should be made. Follow-up time will 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.

4.4 Outcomes/events

AESI (See ACCESS BGR protocol in EU PAS register), identified based on algorithms decided in ACCESS background rate studies that is conducted for EMA. If needed additional AESI may be added.

Definitions, codes, and proposed algorithms for all AESI can be found at https://drive.google.com/drive/folders/1Y_3cuGRN1g-jBv2ec1fC0aYcpxEjtrY9. These AESI definition templates can serve as example for event definitions forms for non-anticipated events.

AESI will be identified in each participating database using diagnosis codes, where possible algorithms should be used to ascertain the event or for sensitivity analyses.

For rapid assessment studies, validation of event identification algorithms using chart abstraction or manual verification of electronic records while being blinded to the Moderna vaccine exposure may be considered. If resources are restricted a sample may be validated initially to assess the positive predictive value. If the PPV is below 80% all cases should be validated.

Certainty of the diagnosis of an event may be classified against the existing and new Brighton Collaboration (BC) case definitions. SPEAC is providing a toolbox to those case definitions which is accessible from the Brighton Collaboration website or by writing to the bc-coordinator@brightoncollaboration.us.

4.5 Exposure assessment

For the ecological design, exposure assessment is based on the date of start of immunization with the Moderna vaccine.

For the SCRI study, exposure to Moderna vaccine should be specific and exposure to other COVID-19 vaccines should be excluded.

Vaccine information should be obtained from all possible sources that capture COVID-19 vaccine and influenza vaccines such as pharmacy dispensing records, general practice records, immunization registers, vaccination records, medical records, or other secondary data sources. Depending on the data source, vaccines may need to be identified via Anatomical Therapeutic Chemical codes, nationally used product codes, batch numbers, local codes, or free text. Data access providers should conduct quality assessment to assess whether vaccinations are adequately captured against external benchmarks using methods as described in ADVANCE system testing (Sturkenboom, 2020).
All doses should be identified

4.6 Covariate Assessment

For descriptive purposes, sex, age, country, and calendar month of vaccination should be assessed in the study population at the start of the study period and at the event date.
Important confounders are all factors that are known risk factors for the event and are associated with

the COVID_19 vaccination.

Ecological Designs (ITS)

Confounding in ecological designs may occur since comparisons are conducted at group level. These are inherent to the design and therefore such analyses will not provide causal evidence. Confounding due to additional interventions concurrent to the intervention of interest should be considered. This may be addressed through inclusion of a control time series if an appropriate control can be identified (Bernal, 2018). This control series may provide a counter-factual, estimating what the trends in incidence of the event of interest would have been in the absence of exposure, given exposure to other relevant time-varying events or interventions. The control series should be composed of a population observed during the same periods, exposed to other changes, but not exposed to the intervention. In the case of healthcare workers targeted in the initial roll-out of COVID-19 vaccines, an appropriate control series may be similarly aged adults who are not healthcare workers.

Self-Controlled Risk Interval Design

Within-person confounding is implicitly controlled for by design. Time-varying confounders will be controlled for in SCRI analyses.

4.7 Study Size

Rapid assessment
In rapid assessment the purpose is to quickly explore whether a potential safety signal can be confirmed, without doing proper adjustments for confounding. The null hypothesis is that the Moderna vaccine does not increase the risk of AESI. For this analysis, all available data that can be rapidly accessed should be utilized. When the null hypothesis cannot be accepted, the rapid assessment confirms the signal but does not show a causal relation as it may still be explained by confounding or bias; when the null hypothesis cannot be rejected, the existence of a true signal cannot be definitively refuted. The upper limit of the confidence interval will show what range of risk is compatible with the data.

5 Data Collection and Management

The work will be conducted using a distributed network of DAPs who agreed to use a common protocol, common data model and common analytics. We will work according to model C (figure 5)

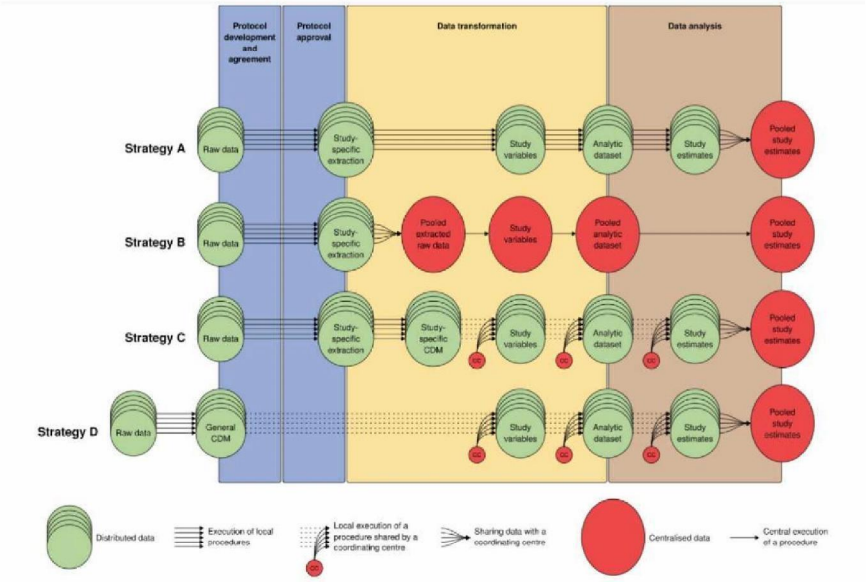

Figure 5. Options for multi-database studies in Europe

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 available[3]. 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.

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 center for pooled meta-analysis and reporting.

Routine procedures will 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 will be followed by each DAP and the coordinating organization, with periodic backups. Standard procedures should be in place at each research center to restore files in the event of a hardware or software failure.

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

6 Data Analysis

6.1 Descriptive Analysis

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

For each study design, demographic characteristics of the study population (e.g. age at study entry, sex) and baseline characteristics (e.g. Co-morbidities) will be summarized for each data source using descriptive statistics where available.

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

Event counts should be provided categorized by level of severity/certainty. Appendices should provide code /algorithm counts for the events.

6.2 Measures of Association

For ecological analysis, incidence rates in different periods will be computed. The ratio between incidence rates in different periods will be computed using a Poisson regression model.
For interrupted time series analyses on incidence rates, a segmented Poisson regression model should be used to estimate changes in level and slope of pre- and post-intervention trends.

For SCRI, the ratio between the incidence rate in the risk period and the incidence rate in the control period (incidence rate ratio) should be computed using conditional Poisson regression.

6.3 Data Integration

Results will be presented separately for each data source and pooled across data sources.
Meta-analysis will be conducted using standard methods: heterogeneity will be tested and Forest plots be provided. Because of the expected variation in effect estimates of data-sources we recommend random effect models (Der Simonian and Laird, 1986).

Subgroup Analysis

If relevant to specific events, the presence of effect modification by relevant variables (age at vaccination, specific comorbidities, concomitant vaccinations) may be assessed using stratification and statistically by testing for interaction.

Sensitivity Analysis

Sensitivity analysis will 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 studies and may include the following:

- For ecological analyses, a transition period after COVID-19 vaccine introduction should be applied to ensure a sufficient vaccine coverage and, consequently ensure to assess effect of vaccine.
- For ecological analyses of a population exposed to targeted vaccination (such as healthcare workers), consider use of a control series in an unexposed population to control for time trends concurrent with but unrelated to the intervention (Bernal, 2018)
- For unadjusted SCRI, if the risk window is not well known, conduct analyses with alternative risk intervals.
- For unadjusted SCRI analyses, 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.

7 Quality Control

Standard operating procedures or internal process guidance at each research center will be adhered to for the conduct of the study. These procedures will 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.

8 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.

9 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 summarized 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-authorization 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 summarized in the interim safety analysis and in the final study report."
Module VIII – Post-Authorization 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.

10 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 authorization holder should communicate to the Agency and the competent authorities of the Member States in which the product is authorized the final manuscript of the article within two weeks after first acceptance for publication."

Protocols will 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 venues will be considered.

11 Other Good Research Practice

This study will adhere to the *Guidelines for GPP* and has been designed in line with the European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP) *Guide on Methodological Standards in Pharmacoepidemiology* (ENCePP, 2018). The *ENCePP Checklist for Study Protocols* (ENCePP, 2018)
The study is a post-authorization study of vaccine safety and will comply with the definition of the non-interventional (observational) study referred to in the International Conference on Harmonization 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-Authorization 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-Authorization 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)

Appendix 1: Description of VAC4EU EHR data sources

Netherlands: PHARMO Database Network

The PHARMO Database Network is a population-based network of electronic healthcare databases and combines anonymous data from different primary and secondary healthcare settings in the Netherlands. These different data sources, including data from general practices, in- and out-patient pharmacies, clinical laboratories, hospitals, the cancer registry, pathology registry and perinatal registry, are linked on a patient level through validated algorithms. To ensure the privacy of the data in the PHARMO Database Network, the collection, processing, linkage and anonymization of the data is performed by STIZON. STIZON is an independent, ISO/IEC 27001 certified foundation, which acts as a Trusted Third Party between the data sources and the PHARMO Institute. The longitudinal nature of the PHARMO Database Network system enables to follow-up more than 9 million persons of a well-defined population in the Netherlands for an average of twelve years. Currently, the PHARMO Database Network covers over 6 million active persons out of 17 million inhabitants of the Netherlands. Data collection period, catchment area and overlap between data sources differ. Therefore, the final cohort size for any study will depend on the data sources included. As data sources are linked on an annual basis, the average lag time of the data is one year. All electronic patient records in the PHARMO Database Network include information on age, sex, socioeconomic status and mortality. Other available information depends on the data source. A detailed description of the different data sources is given below. PHARMO is always seeking new opportunities to link with healthcare databases. Furthermore, it is possible to link additional data collections, such as data from chart reviews, patient-reported outcomes or data from general practice trials.

The General Practitioner database comprises data from electronic patient records registered by GPs. The records include information on diagnoses and symptoms, laboratory test results, referrals to specialists and healthcare product/drug prescriptions. The prescription records include information on type of product, prescription date, strength, dosage regimen, quantity and route of administration. Drug prescriptions are coded according to the WHO ATC Classification System [www.whocc.no]. Diagnoses and symptoms are coded according to the International Classification of Primary Care - ICPC [www.nhg.org], which can be mapped to the International Classification of Diseases - ICD codes, but can also be entered as free text. GP data cover a catchment area representing 3.2 million residents (~20% of the Dutch population).

The Out-patient Pharmacy Database comprises GP or specialist prescribed healthcare products dispensed by the out-patient pharmacy. The dispensing records include information on type of product, date, strength, dosage regimen, quantity, route of administration, prescriber specialty and costs. Drug dispensing are coded according to the WHO ATC Classification System. Out-patient pharmacy data cover a catchment area representing 4.2 million residents (~25% of the Dutch population). PHARMO is listed under the ENCePP resources database. PHARMO data capture influenza vaccine and may be linked to the PRAEVENTIS database that is held by RIVM, based on specific permissions.

Denmark: Danish Registries

Denmark has a tax-funded health care system ensuring easy and equal access to health care for all its citizens, and with this system all contacts are recorded in administrative and medical registers (Schmidt et al., 2019). The records carry a unique personal identification number, called

the CPR-number, assigned to every Danish citizen. Linkage between registers at an individual level is possible because this CPR-number is used in all Danish registers and assigned by the Danish Civil Registration System (Schmidt et al., 2014). All registers have a nationwide coverage and an almost 100% capture of contacts covering information on currently 5.8 million inhabitants plus historical information. For the purpose of the study we will obtain information from the following registries. The Danish National Prescription Registry (DNPR) includes data on all outpatient dispensing of medications and vaccines at Danish pharmacies from 1995 and onwards, including dispensing date, ATC code, product code and amount (Pottegard et al., 2017). The Danish National Health Service Register includes data on primary care services, including general practitioner contacts, examinations, procedures, and vaccinations; psychologist or psychiatrist and other primary care provider visits; etc. From the Danish Civil Registration System, data on demographics (sex, date of birth) and censoring (migration, vital status). The Danish National Patient Registry contains diagnoses and procedures from all hospitalizations since 1977 and contacts to hospital outpatient clinics since 1995 (Schmidt et al., 2015). The Danish National Health Service Register contains information on referral for vaccine administration from GPs (Sahl Andersen et al., 2011). The Danish databases were characterized in the ADVANCE project and considered fit for purpose for vaccine coverage, benefits and risk assessment and could participate in near real-time monitoring (Sturkenboom et al., 2020; Bollaerts et al., 2019).

Spain: SIDIAP

The Information System for Research in Primary Care (Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària' - SIDIAP; www.sidiap.org) was created in 2010 by the Catalan Health Institute (CHI) and the IDIAPJGol Institute. It includes information collected since 01 January 2006 during routine visits at 278 primary care centers pertaining to the CHI in Catalonia (North-East Spain) with 3,414 participating GPs. SIDIAP has pseudo-anonymized records for 5.7 million people (80% of the Catalan population) being highly representative of the Catalan population.

The SIDIAP data comprises the clinical and referral events registered by primary care health professionals (GPs, pediatricians and nurses) and administrative staff in electronic medical records, comprehensive demographic information, community pharmacy invoicing data, specialist referrals and primary care laboratory test results. It can also be linked to other data sources, such as the hospital discharge database, on a project by project basis. Health professionals gather this information using ICD-10 codes, ATC codes and structured forms designed for the collection of variables relevant for primary care clinical management, such as country of origin, sex, age, height, weight, body mass index, tobacco and alcohol use, blood pressure measurements, blood and urine test results. In relation to vaccines, SIDIAP includes all routine childhood and adult immunizations, including the antigen and the number of administered doses. Encoding personal and clinic identifiers ensures the confidentiality of the information in the SIDIAP database. The SIDIAP database is updated annually at each start of the year.

But nowadays, with the COVID-19 pandemic, there is the possibility to have shorter term updates in order to monitor the evolution of the pandemic. Recent reports have shown the SIDIAP data to be useful for epidemiological research. SIDIAP is listed under the ENCePP resources database www.encepp.eu/encepp/resourcesDatabase.jsp). The SIDIAP database was characterized in the ADVANCE project and considered fit for purpose for vaccine coverage, benefits and risk assessment (Sturkenboom et al., 2020).

Spain: FISABIO

The region of Valencia, with 5 million inhabitants, is part of the Spanish National Health System, a universal public healthcare system. Information will be obtained from the population-based electronic information systems of the Valencia Health Agency (VHA) and the regional Government of Valencia: (1) The Population Information System (SIP) provides an identification number for each person under Valencian Health Service (VHS) coverage, and registers some demographic characteristics, and dates and causes of VHA discharge, including death. (2) The minimum basic dataset at hospital discharge is a synopsis of clinical and administrative information on all hospital discharges, including diagnoses and procedures (all electronic health systems in the VHS use the ICD- 9-CM). (3) The Emergency Department module (ED) including ED dates of visit and discharge and reason for discharge. (4) The electronic medical record (EMR) for ambulatory care, available in all primary healthcare centers and other ambulatory settings. It has all the information on patients regarding diagnoses, their personal and family medical history, laboratory results, lifestyle, etc. (5) The pharmaceutical module (prescription information system), part of EMR, includes information about both physician prescriptions and dispensations from pharmacy claims. (6) The Corporate Resource Catalogue provides information about the geographical and functional organization of VHS, its health centers, health services provided and professionals in healthcare. Specific public health registries are available and linkable at an individual level (such as the perinatal registry and the congenital anomalies registry, from which pregnancy outcomes can be obtained) All the information in these systems can be linked at an individual level through the SIP number. The FISABIO database was not characterized in ADVANCE, but did provide important information evaluating the pandemic influenza vaccine and narcolepsy (Dodd CN et al., 2018; Weibel et al., 2019)

Spain: Valencia Integrated Databases (VID)

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

Italy: PEDIANET database

PEDIANET, a pediatric general practice research database, contains reason for accessing healthcare, health status (according to the Guidelines of Health Supervision of the American

Academy of Pediatrics), demographic data, diagnosis and clinical details (free text or coded using the ICD-9 CM), prescriptions (pharmaceutical prescriptions identified by the ATC code), specialist appointments, diagnostic procedures, hospital admissions, growth parameters and outcome data of the children habitually seen by about 140 family pediatricians (FPs) distributed throughout Italy.

PEDIANET can link to other databases using unique patient identifiers. In the first database, information on routine childhood vaccination are captured including vaccine brand and dose. In the second database, information on patient hospitalization date, reason for hospitalization, days of hospitalizations and discharge diagnosis (up to six diagnosis) are captured. The FPs participation in the database is voluntary and patients and their parents provide consent for use of their data for research purposes. In Italy each child is assigned to a FP, who is the referral for any health visit or any drug prescription, thus the database contains a very detailed personal medical history. The data, generated during routine practice care using common software (JuniorBit®), are anonymized and sent monthly to a centralized database in Padua for validation. The PEDIANET database can be linked to regional vaccination data which was successfully tested in the ADVANCE project where it was characterized and deemed fit for purpose for pediatric routine vaccines (Sturkenboom et al., 2020).

Italy: ARS database

The Italian National Healthcare System is organized at regional level: the national government sets standards of assistance and a tax-based funding for each region, and regional governments are responsible to provide to all their inhabitants. Tuscany is an Italian region, with around 3.6 million inhabitants. The Agenzia Regionale di Sanita' della Toscana (ARS) is a research institute of the Tuscany Region. The ARS database comprises all information that are collected by the Tuscany Region to account for the healthcare delivered to its inhabitants. Moreover, ARS collects data from regional initiatives. All the data in the ARS data source can be linked with each other at the individual level, through a pseudo-anonymous identifier. The ARS database routinely collects primary care and secondary care prescriptions of drugs for outpatient use, and is able to link them at the individual level with hospital admissions, admissions to emergency care, records of exemptions from copayment, diagnostic tests and procedures, causes of death, mental health services registry, birth registry, spontaneous abortion registry, induced terminations registry. A pathology registry is available, mostly recorded in free text, but with morphology and topographic Snomed codes. Mother-child linkage is possible through the birth registry. Vaccine data is available since 2016 for children and since 2019 for adults. However, to date, 2019 vaccination data for adults may still be incomplete. The ARS database was characterized in the ADVANCE project and considered fit for purpose for vaccine coverage, benefits and risk assessment when using the new vaccine registry (from 2019) (Sturkenboom et al., 2020).

Italy: Lazio database

The Italian National Healthcare System is organized at regional level: the national government sets standards of assistance and a tax-based funding for each region, and regional governments are responsible to provide to all their inhabitants. Tuscany is an Italian region, with around 3.6 million inhabitants. The Lazio database comprises all information The Lazio regionale database routinely collects primary care and secondary care prescriptions of drugs for outpatient use, and is able to link them at the individual level with hospital admissions, admissions to emergency care, records of exemptions from co-payment, diagnostic tests and procedures, causes of death, mental health services registry, birth registry, spontaneous abortion registry, induced terminations registry. A pathology registry is available, mostly recorded in free text, but with

morphology and topographic Snomed codes. Mother-child linkage is possible through the birth registry.

Italy: HSD database

The Health Search CSD Longitudinal Patient Database (HSD), is a longitudinal observational database that is representative of the Italian general population. It was established in 1998 by the Italian College of General Practitioners. The HSD contains data from computer-based patient records from a selected group of GPs (covering a total of 2 million patients) located throughout Italy. These GPs voluntarily agreed to collect data for the database and attend specified training courses. The database includes information on the age, gender, and identification of the patient, and GP registration information, which is linked to prescription information, clinical events and diagnoses, hospital admission, and causes of death. Laboratory values are available. All diagnoses are coded according to the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). Drug names are coded according to the ATC classification system. To be included in the study, GPs must have provided data for at least 1 year and meet standard quality criteria pertaining to: levels of coding, prevalence of well-known diseases, and mortality rates. The HSD complies with EU, guidelines on the use of medical data for research. The HSD has been used as data source for a number of peer reviewed publications on the prevalence of disease conditions, drug safety and prescription patterns in Italian primary care. Approval for use of data is obtained from the Italian College of General Practitioners.
HSD is listed under the ENCePP resources database (www.encepp.eu/encepp/resourcesDatabase.jsp).

United Kingdom: CPRD & HES

The Clinical Practice Research Datalink (CPRD) from the UK collates the computerized medical records of general practitioners (GPs) in the UK who act as the gatekeepers of healthcare and maintain patients' life-long electronic health records. As such they are responsible for primary healthcare and specialist referrals, and they also store information stemming from specialist referrals, and hospitalizations. GPs act as the first point of contact for any non-emergency health-related issues, which may then be managed within primary care and/or referred to secondary care as necessary. Secondary care teams also feedback information to GPs about their patients, including key diagnoses. The data recorded in the CPRD include demographic information, prescription details, clinical events, preventive care, specialist referrals, hospital admissions, and major outcomes, including death. The majority of the data are coded in Read Codes. Validation of data with original records (specialist letters) is also available.
The dataset is generalizable to the UK population based upon age, sex, socioeconomic class and national geographic coverage when GOLD & Aurum versions are used.
There are currently approximately 42 million patients (acceptable for research purposes) – of which 13 million are active (still alive and registered with the GP practice) – in approximately 1,700 practices (https://cprd.com/Data). Data include demographics, all GP/healthcare professional consultations (phone, letter, email, in surgery, at home), diagnoses and symptoms, laboratory test results, treatments, including all prescriptions, all data referrals to other care, hospital discharge summary (date and Read codes), hospital clinic summary, preventive treatment and immunizations, death (date and cause). For a proportion of the CPRD panel practices (>80%), the GPs have agreed to permit CPRD to link at patient level to the Hospital Episode Statistics (HES) data. CPRD is listed under the ENCePP resources database, access will be

provided by the Utrecht University. The CPRD was not yet characterized in the ADVANCE project, where the UK THIN and RCGP databases were used, but has been largely used in vaccine studies.

The HES database contains details of all admissions to National Health System (NHS) hospitals in England; approximately 60% of GP practices in the CPRD are linked to the HES database. Not all patients in the CPRD have linked data (e.g. if they live outside England or if their GP has not agreed that their data should be used in this way). As with standard CPRD patients, HES data are limited to research-standard patients. CPRD records are linked to the HES using a combination of the patient's NHS number, gender and date of birth (**Error! Reference source not found.** et al., 2012).

France: Système National des Données de Santé (SNDS)

The SNDS (Système National des Données de Santé) is the French nationwide healthcare database. It currently covers the overall French population (about 67 million persons) from birth (or immigration) to death (or emigration), even if a subject changes occupation or retires. Using a unique pseudonymized identifier, the SNDS merges all reimbursed outpatient claims from all French health care insurance schemes (SNIIRAM database), hospital-discharge summaries from French public and private hospitals (PMSI database), and the national death registry. SNDS data are available since 2006 and contains information on:

- General characteristics: gender, year of birth, area of residence, etc.
- Death: month, year and cause
- Long-Term Disease registration associated with an ICD-10 diagnostic codes
- Outpatient reimbursed healthcare expenditures with dates and codes (but not the medical indication nor result): visits, medical procedures, nursing acts, physiotherapy, lab tests, dispensed drugs and medical devices, etc. For each expenditure, associated costs, prescriber and caregiver information (specialty, private/public practice) and the corresponding dates are provided.
- Inpatients details: primary, related and associated ICD-10 diagnostic codes resulting from hospital discharge summaries with the date and duration of the hospital stay, the performed medical procedures, and the related costs. Drugs included in the diagnosis related group cost are not captured. However, expansive drugs (i.e. the one charged in addition to the group cost) are.

Outpatient data (SNIIRAM) are uploaded to the SNDS throughout the year. It is admitted that a lag of around 6 months is required to catch 90% of the dispensings. Inpatient data (PMSI) are uploaded in one time, at the end of the following year. Hence, we consider that complete SNDS data of year Y are available in January of the year Y+2. SNDS access is regulated.
Each study and data extraction need approval from the CESREES (*Comité Ethique et Scientifique pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé*) in charge of assessing scientific quality of the project, and authorization from the CNIL (French data protection commission), and then contracts with the SNDS data holder (CNAM) for data extraction. Bordeaux PharmacoEpi (BPE), a research platform of the University of Bordeaux specialized in real world studies, will be in charge of requesting access to SNDS data. The SNIIRAM data were not yet characterized in the ADVANCE project but have been used for vaccine studies.

Norway

The core data that UIO has access to is the health care administrative databases of the entire Norwegian population, which amounts to approximately 5.3 million inhabitants.

Norway has a universal public health care system, consisting of primary health care services and specialist health care services. Many population-based health registries were established in the 1960s, with use of unique personal identifiers facilitating linkage between registries. The mandatory national health registries were established to maintain national functions. They are used for health analysis, health statistics, improving the quality of healthcare, research, administration and emergency preparedness. The most commonly used registries are administrated by The Norwegian Institute of Public Health, The Norwegian Directorate of Health and Statistics Norway. Information about all Norwegian National Registries can be found here: www.fhi.no/en/more/access-to-data/about-the-national-health-registries2/

The Norwegian national identity number was introduced more than 50 years ago. This identifier is assigned to every person at birth or upon immigration; it is 11 digits long and encodes date of birth and gender. The code is included in all national registries, allowing accurate linkage among them.

Germany: GePaRD

GePaRD is based on claims data from four statutory health insurance providers in Germany and currently includes information on approximately 25 million persons who have been insured with one of the participating providers since 2004 or later. Per data year, there is information on approximately 20% of the general population and all geographical regions of Germany are represented. In addition to demographic data, GePaRD contains information on dispensations of reimbursable prescription drugs as well as outpatient (i.e., from general practitioners and specialists) and inpatient services and diagnoses. GePaRD also contains information on influenza vaccinations and routine childhood immunizations and there is experience with studies on utilization and risk of vaccination and on background incidence of adverse events of vaccinations (Hense et al., 2014; Schink et al., 2014). GePaRD data have been used for vaccine safety studies. GePaRD is listed under the ENCePP resources database.