Innovative Medicines Initiative

ADVANCE

www.advance-vaccines.eu

Accelerated Development of VAccine beNefit-risk Collaboration in Europe

Grant Agreement nº115557

D5.9 White paper

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

V1.0

Lead beneficiary: P-95
Date: February 2, 2018
Nature: Report
Dissemination level: PU

DOCUMENT INFORMATION

Grant Agreement Number	115557	**Acronym**	ADVANCE
Full title	Accelerated Development of VAccine beNefit-risk Collaboration in Europe		
Project URL	http://www.advance-vaccines.eu		
IMI Project officer	Angela Wittelsberger (angela.wittelsberger@imi.europa.eu)		

Deliverable	**Number**	5.9	**Title**	White paper
Work package	**Number**	5	**Title**	Proof-of-concept studies of a framework to perform vaccine benefit-risk monitoring

Delivery date	**Contractual**	Month 48	**Actual 51**	Month
Status	Draft version / V1.0		Draft ☐ Final x	
Nature	Report ☒ Prototype ☐ Other ☐			
Dissemination Level	Public x Confidential			

Authors (Partner)	Miriam Sturkenboom, Lina Titievsky			
Responsible Author	Miriam Sturkenboom		**Email**	miriam.sturkenboom@p-95.com
	Partner	P-95	**Phone**	+31-657831983

DOCUMENT HISTORY

NAME	DATE	VERSION	DESCRIPTION
Miriam Sturkenboom	3-11-2017	0.1	First draft
Lina Titievsky	6-11-2017		Comments
Miriam Sturkenboom	7-11-2017	0.2	Updated and extended for first SC check
Miriam Sturkenboom	16-11-2017	0.3	Update after SC feedback
Lina Titievsky	19-11-2017		Review, edits and comments, examples
Miriam Sturkenboom	22-11-2017	0.4	Trimming and addressing comments
Miriam Sturkenboom	23-11-2017	0.5	Addition of executive summary and annex 2
Vincent Bauchau, Alena Khromava, Linda Levesque, Sonja Banga, Simon de Lusignan, Kaat Bollaerts, Mendel Haag, Laurence Pagnon, Tyra Grove Krause, Gino Picelli, Tin Tin Htar Myint, Patrick Mahy, Consuelo Huerta, Elisa Martin, Daniel Weibel, Hanne Dorthe Meborg, Jorgen Bauwens, Boni Bolibar, Xavier Kurz, Rosa Gini, Toon Braeye, Matti Lehtinen, Chris Mc Gee, Lieke van der Aa	8-12-2017		Review of deliverable and comments
Lina Titievsky, Miriam Sturkenboom	29-12-2017	0.6	Extensive edits via addressing comments from SC & WP5
Lieke van der Aa, Patrick Mahy, Vincent Bauchau, Xavier Kurz	10-1-2018		Comments and discussion CoI management
Miriam Sturkenboom	31-1-2018	1.0	Address comments and update annex 3

Table of Contents

DOCUMENT INFORMATION 2

DOCUMENT HISTORY 3

DEFINITIONS 5

1. Executive summary 8

1. Context 12

2. Which processes, methods and systems were tested? 13

2.1 Testing inclusivity of all stakeholders 14

2.2 Testing the ADVANCE Code of Conduct & Practice Guidance on Quality Management 15

2.3 Testing whether data sources are 'fit for purpose' to participate in vaccine coverage, benefits and risk studies 17

2.4 Testing the methods to generate data on vaccination coverage, benefits and risks 22

2.5 Testing the platform and work processes to generate evidence 24

2.6 POC Evaluation 27

3 The future and next steps 27

Annex 1: Recommendation in the ADVANCE CoC and implementation in the POC 29

Annex 2 : POC Implementation of quality framework as queried by WP1 32

Annex 3: Aggregated fingeprint data provided to the remote resource environment from different data sources 35

Annex 4 ADVANCE Process map 36

DEFINITIONS

Partners 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)
- **ASLCR.** Azienda Sanitaria Locale della Provincia di Cremona (Italy)
- **AUH.** Aarhus Universitetshospital (Denmark)
- **J&J.** Janssen Vaccines - Prevention B.V. (Belgium)
- **ECDC.** European 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.** Royal College of General Practitioners (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)
- **WIV-ISP.** Belgian Scientific Institute of Public Health (Belgium)

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 f**oundation (Switzerland)
- **WKT** State Medicines Control Agency (Lituania)
- **WUM** Polish Medicines Agency (Poland)

This preliminary glossary of terms used in the review will be further developed in ADVANCE accumulating the contributions for the various deliverables and work packages.

- **ADVANCE:** Accelerated Development of VAccine beNefit-risk Collaboration
- **Aggregated Data:** The cumulative or summary information that does not specifically identify any particular person[1]
- **AIRR:** ADVANCE International Research Readiness Instrument
- **B/R**: benefit-risk
- **CoC:** Code of Conduct
- **DAP: data access provider**
- **EMIF:**
- **IMEDS** Innovation in Medical Evidence Development and Surveillance
- **CNODES** Canadian network for Observational Drug Effect Studies
- **CRO:** Clinical Research Organization
- **(Datasource) Fingerprinting:** characterization of the actual data characteristics of a database.
- **ETL:** Extract, Transform, Load
- **EU PAS:** European Union electronic Register of Post-Authorisation Studies
- **GDPR**: General Data Protection Regulation
- **Harmonised Data:** The harmonised data follow a consensus and are formatted in the same way

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

- **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/her physical, physiological, mental, economic, cultural, or social identity."
- **Ontology:** An ontology is defined as a hierarchy, or specification of clinical concepts and their relationships within a given domain[2]
- **Original Data:** Data, as maintained by the Data Source or any organization which collects the data, before inclusion in the platform
- **Proof-of-concept studies (POC):** studies that will be conducted during ADVANCE project to evaluate and provide evidence of whether proposed methods, guidance documents and methods of integration and collaboration work
- **RRE**: Remote research environment
- **VSD**: Vaccine Safety Datalink

[2] Gruber T. A translation approach to portable ontology specifications. Knowledge Acquisition 1993; 5 (2):199–220. doi:10.1006/knac.1993.1008

1. Executive summary

This white paper describes the context and lessons learned in testing the Accelerated Development of VAccine beNefit-risk Collaboration (ADVANCE) prototype distributed network system (inclusivity of public and private stakeholders, code of conduct, methods, eligibility of data sources, data sharing workflows/ tools and stakeholder feedback) and the recommendations following these.

The ADVANCE system to generate evidence on the benefits and risks of vaccines from existing health data was tested by: conducting a fit for purpose assessment of different electronic health care data sources accessible to ADVANCE partners, and four protocol based studies (comprising the proof of concept (POC)) on coverage, benefits, risks and benefit-risk integration using as example pertussis vaccination.

We list the key lessons learned per tested area:

Inclusivity (public-private collaboration)

1) Four studies were completed successfully involving 68 study team members from all stakeholders (academic, private research organizations, vaccine manufacturers, regulators and public health institutes).
   Public-private collaboration was possible following the ADVANCE code of conduct - Reaching consensus across all stakeholders required substantial time.

ADVANCE Code of Conduct & quality guidance

1) Procedures to implement the 10 topics of the ADVANCE Code of Conduct (CoC) were successfully implemented and adhered to in the four protocol based studies, each of which was registered in the European Union electronic Register of Post-Authorisation Studies (EU PAS register).
2) The concepts in the Good Practice Guidance on Quality Management (i.e. epidemiologic /scientific best practices ranging across key steps and/or deliverables covering protocol development, analysis, data access, ethics, compliance, etc) were mostly aligned but not formally assessed with the proposed criteria on Quality Management for ADVANCE

ADVANCE innovative methods

1) Novel methods for dealing with heterogeneity across electronic health data sources, validity of the outcomes, and integration of benefits and risks of vaccines using either a Multi Criteria Decision Analysis or a visualization tool (i.e. a dashboard) were assessed and/or implemented using real world data and demonstrated potential value.
   Visualization dashboards depicting real world coverage, benefits and risk data were appreciated and may present a good alternative for monitoring benefits and risk over time.

ADVANCE data sources

1) Nineteen diverse data sources (general practice databases, record linkage, disease surveillance, cohorts) from 8 countries (Finland, Sweden, Denmark, UK, Netherlands, Spain, Italy, Belgium) were ultimately accessible to ADVANCE partners. Based on these 19:
    a. Fourteen (74%) data access providers (DAP) provided descriptive characterization data content on the remote research environment (RRE) by November 2017,
    b. 10 (53%) DAPs provided the output of the population characterization scripts and incidence rates of up to 25 events of interest, covering 50 million persons.
    c. Ten of Daps provided coverage data for pertussis, measles, HPV and/or influenza vaccination based on R scripts. Two infectious disease surveillance systems (NL, BE) provided data on case counts.
2) Efficient, rigorous and transparent processes and procedures akin to standard operating procedures were implemented for assessment whether data sources are 'fit for purpose' for specific studies based on meta-data from the ADVANCE International Research Readiness Instrument (AIRR) survey, standardized characterization of actual data content and comparison with external and internal benchmarks
3) Seven out of 19 databases (covering 35 million persons source population, in 4 countries) were ready and eligible in 2016 to participate in the POC on pertussis vaccination and completed these successfully.

ADVANCE data sharing platform

1) ADVANCE successfully implemented a distributed data network system, comparable to the Vaccine Safety Datalink (VSD), Sentinel and Canadian network for Observational Drug Effect Studies (CNODES) distributed networks, on a limited budget and introduced a quality assessment and conducted four protocol based studies on pertussis vaccine coverage, benefits, risks and benefit/risk analysis.
2) For efficiency ADVANCE used a project-based simple common data model for population, events and vaccinations.
3) The ADVANCE Code-mapper open source web-application tool proved to be very useful for semantic harmonization of events across different disease dictionaries.
4) Vaccination cleaning and harmonization were novel and most laborious for data sources; details on brands were available only in few data sources.
5) The newly developed ADVANCE multilingual Vaccine Ontology (VaccO) may support data access providers with semantic harmonization of vaccines in the future and further labelling of vaccines (e.g. brand).
6) Data transformation of common input files was harmonized and conducted with double coded Jerboa or R/SAS scripts

7) Aggregated data (covering populations of more than 50 million persons) were shared by all participating DAPs on a secure central platform for further analysis & pooling
8) Capacity and readiness to conduct vaccine studies using a distributed network like model was built amongst 68 persons in the study teams including vaccinologists, statisticians and data access providers.

Evaluation

1) Based on the analysis of steps in the proof of concept studies, a detailed process map for conduct of protocol based studies was produced, facilitating planning of the future studies with increased efficiencies.
2) Internal stakeholders feedback from ADVANCE partners showed high appreciation of the collaboration, processes, tools and transparency, but also a need to provide more clarity, risk management of declarations of interest and to improve on speed and ability to address novel scientific vaccine questions.
3) Formal feedback from an external panels (regulatory agencies, WHO, academics and public health) organized by ECDC on the POC revealed that the results and methods are appreciated and promising with more necessary work to be done in the future : "added value in practice is not yet fully clear, for reasons including doubts over the quality, representativeness and comparability of data derived from different types of medical electronic databases (as compared with other sources); lack of validation through chart review; the small number of countries involved and limited population coverage; the lack of vaccine product-specific information; and the length of time needed to access data and generate results. Nevertheless, the POC results show potential, and sustainability of the platform should be assured with a view to continual improvements in quality, validation and timeliness until the results is at a level where they can be used to protect public health".

Recommendations

- We recommend disseminating widely the available ADVANCE guidance documents and proposed procedures, methods and tools through publicly available webinars and/or e-learning tools.
- We recommend publishing the methods and findings in peer reviewed journals for wider feedback, adaptability and scientific credibility of the ADVANCE approach.
- We recommend that readiness to act and generate high quality evidence quickly by data access providers, statisticians and epidemiologists across, member states should be fostered and resourced adequately on a national level.
- We recommend that semantic harmonization of more events and more vaccines and transformation into a common data model is continued to increase data readiness and speed.

- Based on the proven ability to generate evidence on vaccine benefits and risks from available health data throughout Europe in a true public private collaboration, we recommend sustaining and extending this.

In conclusion, based on 1) the successful implementation of fit for purpose assessment of different electronic health care data sources in 12 data sources; 2) a distributed data network system to generate evidence on vaccine coverage, benefit and risk studies across 8 member states with secondary use of health data in a public-private collaboration and 3) the feedback obtained, it appears to be feasible to increase speed, scale (geographically and population wise), details on vaccines through the ontology and quality of outcomes, and subsequent transformation of the current ADVANCE model into a more sustainable production phase.

1. Context

Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is an ongoing European public-private collaboration that was initiated in 2013 and is scheduled to end in 2018. It was funded by the Innovative Medicines Initiative (IMI), a joint undertaking by the European Union (EU) and European Pharmaceutical industry (EFPIA).

ADVANCE was created in response to the 2009 (H1N1) flu pandemic. The European experience highlighted the factors limiting the capacity to collect European data on vaccine exposure, safety and effectiveness, including[3]:

- Lack of rapid access to available data sources or expertise,
- Difficulties in establishing efficient interactions between multiple stakeholders,
- Concerns about possible or actual conflicts of interest (or perceptions thereof), and
- Inadequate public funding to generate the required benefit and risk data and inability of private partners to collaborate with public health institutes to generate the regulatory required data

Consequently, ADVANCE seeks to address the feasibility of establishing a public-private collaboration to respond to relevant public health questions associated regarding the benefits and risks of vaccines in a timely and efficient manner.

The ADVANCE vision is to deliver "*Best evidence at the right time to support decision-making on vaccination in Europe*", and its mission is to establish a prototype of a sustainable and compelling system that rapidly provides best available scientific evidence on post-marketing vaccination benefits and risks for well informed decisions. Therefore, ADVANCE involves the creation and assessment of an infrastructure (i.e. system) that would bring together different stakeholders and data sources in Europe.

The benefits and risks of vaccines are perceived and weighed differently, and at different times from other medicinal products as they are often offered prophylactically to healthy individuals' e.g. as part of the national childhood vaccination programs. As such, vaccines have major public health implications. In addition, vaccines get a lot of media attention. Hence, the tolerance for risk, even if it is an easily treated condition, is very low, as current debates in several EU member states demonstrate. Stakeholders working in the vaccine area therefore need to monitor relevant data continuously and need to have data easily available for quick decision making and risk mitigation.

In the ADVANCE concept, evidence on vaccine coverage, benefits and risks is to be generated through *secondary use of existing health care data in Europe*. We tested this concept by working closely with ADVANCE partners who have access to variable data sources including general practice databases, claims databases, vaccine registries, vaccine trial cohorts and disease surveillance data, resulting in heterogeneity of data. The aim was to test whether the ADVANCE system could permit the rapid

[3] Sturkenboom M. Advancing collaborative vaccine benefits and safety research in Europe via the ADVANCE code of conduct. Vaccine. 2018 Jan 4;36(2):194-195.

generation of information on benefits, coverage and risks of vaccines from these data sources both in the characterization as well as in the conduct of specific studies. In order to maximally take advantage of these different data, ADVANCE has implemented a distributed network model comparable to existing networks in the US (Sentinel, Vaccine Safety Datalink) and Canada (CNODES), although differences exist between the different approaches.

This White Paper describes the context, lessons learned during implementation and testing of the ADVANCE system as well as providing recommendations for the future.

2. Which processes, methods and systems were tested?

The testing of the ADVANCE system was performed by executing four different protocol based studies (also referred to as four pillars), comprising the proof of concept study (POC), and using pertussis vaccines as an example. This vaccine was chosen after collecting input on potential studies from the consortium and ranking of proposals by the Steering Committee. Pertussis vaccine ranked high on the list of indicators because much is known (i.e. published literature and extensive post-marketing experience) about this vaccine which is helpful for benchmarking; furthermore, it continues to be of public health relevance as it is highly utilized and multiple companies/partners have a pertussis vaccine (we did not want to single out one manufacturer). In these studies, we analyzed *retrospectively* the coverage, benefits and risks of the pertussis vaccines in multiple EU countries.

More specifically, we performed the POC to test the ability of the system to:

1) Include important stakeholders ultimately integral to the benefit/risk evaluation of a given vaccine in study teams (public health institutions, regulatory authorities, public and private research organizations (CROs, vaccine manufacturers);
2) Implement the ADVANCE Code of Conduct;
3) Set-up a network of 'fit for purpose' databases able to participate in vaccine coverage, benefits and risks studies;
4) Implement new methods designed to generate rapidly valid evidence on vaccination coverage, benefits and risks and improve the scientific credibility of the studies performed through the system;
5) Share data securely using the ADVANCE workflows and a remote research environment (server behind firewalls, that allows secure data uploading and remote access to analyze)[4]. Details about the infrastructure are available in D5.1 and D5.6.
Our recommendations are based on the lessons learned by the operations team during the conduct of the studies, the formal evaluation of the studies via surveys and workshops by the POC evaluation team (i.e. a specific study team that was created for the evaluation of the POC throughout the scientific and

[4] Trifirò G, Coloma PM, Rijnbeek PR, Romio S, Mosseveld B, Weibel D, Bonhoeffer J, Schuemie M, van der Lei J, Sturkenboom M. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? J Intern Med. 2014 Jun;275(6):551-61.

operational journey), and the feedback from the ECDC Implementability Advisory Board (i.e. an external group of individuals coordinated by the ECDC that reviewed and provided feedback on ADVANCE deliverables).

2.1 Testing inclusivity of all stakeholders

Background/context

The proof of concept was designed to be *inclusive* in membership and conducted as part of the IMI public-private partnership, where private commercial and non-commercial organizations (vaccine manufacturers & private research institutions and foundations), public health institutions, regulatory agencies, and public research organizations were **equal** partners in decision making and study teams.

Learning

The four studies comprising the POC demonstrated that it is possible to work harmoniously together as equal partners in study teams and find consensus. Multi-stakeholder study teams, involving 68 persons in total, demonstrated extensive, open collaboration in generating evidence on the benefits and risks of vaccines. However, reaching an agreement and maintaining trust when the impact of the decision differs largely between stakeholders, e.g. a choice of the vaccine to be studied, regulatory status associated with the studies underlying the POC required appreciable time. The key enabler was to acknowledge the different needs/perspectives during the decision-making processes for the studies. While all stakeholders require that trusted and scientifically rigorous evidence is obtained rapidly while putting the public interest first, understanding of individual objectives and the context within which partners and stakeholder work is imperative as highlighted below.

Vaccine manufacturers need to work in a highly regulated environment, including the obligation to monitor the benefit-risk balance of vaccines during their life cycle. When studies are conducted, manufacturers need to follow company standard operating procedures that reflect regulatory obligations and scientific rigor. They have extensive experience and expertise in generating scientific evidence on vaccines and vaccination. They need to produce specific information in a timely fashion if issues arise. Having *trusted evidence rapidly* in the best way possible is of major interest.

Public health institutes: often have access to relevant data in the context of vaccine studies/research questions (e.g. infectious disease surveillance, vaccination registries) and are experienced in generating the evidence needed to support their vaccination programs. They need to inform Ministries of Health and the public in a timely fashion if an issue arises. We learned that having *trusted evidence rapidly* in the best way possible is their main interest, and that reaction time and public trust are keys for them.

Regulatory agencies: have the responsibility to assess the benefit-risk balance of a licensed vaccine during its lifecycle. Increasingly, regulators use a wide range of evidence from various complimentary sources to support decision-making, including evidence generated independently by pharmaceutical companies. Getting rapid access to independent and valid product specific *scientific evidence* is a major priority.

Clinical Research Organization (CROs): provide fee-for-service research services for clients, who may be private or public entities depending on a research questions/context, in a competitive environment. They have to comply with regulatory obligations while respecting the client's procedures and timelines.

Academic research institutes: academics focus on innovation, multidisciplinary research and teaching. They can provide important expertise in the development of methods and tools, and scientific credibility and innovation are their main priorities. They may also provide contracted research projects for vaccine manufacturers within the framework of the academic code of conduct, e.g. freedom of publication.

Data access providers (DAPs). Access to data was provided by some of the above-mentioned partners. Along with the constraints and needs linked to the nature of the organization (as described above), the access to data comes to those partners with additional obligations. DAPs have privacy regulations and other ethical and legal constraints to comply with. More in general, they must be accountable to the public and individual for the use of the data they are allowed to access to and reputation is a key driver for them. Indeed, in the majority of cases, DAPs do not *own* the data, and conditions for access may change based on public perception.

Recommendations
We recommend that public-private collaborative studies be conducted with all types of stakeholders, since there is a clear added value in terms of knowledge and capacity; this will be easier with long-lasting collaborations and increased trust. However, safeguards need to be in place in case of potential conflicts of interest, and the ADVANCE Code of Conduct should be followed in all studies. The structure for the future governance should facilitate rapid decision making when timelines are tight.

2.2 Testing the ADVANCE Code of Conduct & Practice Guidance on Quality Management

Background/context
The ADVANCE Code of Conduct (CoC) was written to provide a code of conduct for the execution of studies in public-private collaborations and includes 45 recommendations for 10 topics[5]
We operationalized this guidance document by creating simple ADVANCE operating study procedures and workflows applicable to study teams and data access providers (details in D5.3 & D5.6). Annex 1 summarizes the details per CoC topic.
A Good Practice Guidance on Quality management for ADVANCE was described in D1.9[6] (see annex 2) and incorporates the epidemiologic/scientific best practices ranging across key steps and/or deliverables. At this time, we used this guidance for an informal self-assessment.

[5] Kurz X, Bauchau V, Mahy P, Glismann S, van der Aa LM, Simondon F; ADVANCE consortium. The ADVANCE Code of Conduct for collaborative vaccine studies. Vaccine. 2017 Apr 4;35(15):1844-1855.
[6] http://www.advance-vaccines.eu/app/archivos/publicacion/16/ADVANCE_WP1_Deliverable-1.9_Final-Public.pdf

Learning

The ADVANCE CoC recommendations on such topics as scientific integrity, scientific independence, transparency, conflicts of interest, study protocol and study report could all be successfully implemented as demonstrated in the POC. Internal ADVANCE stakeholder feedback obtained at the General Assembly Meeting in Barcelona in September 2017 demonstrated that stakeholders appreciated the transparency of all the processes, documents including having documented declarations of interest. Although this transparency is necessary, it is insufficient to provide clarity (e.g. a clear summary and overview). We also learned that stakeholders wish to have a more explicit conflict of interest management process, especially for the post-ADVANCE stage where multiple (e.g. academic, public health, financial) and real interests may be at stake. The topic "research contract" could not be tested since all studies executed were part of the IMI-ADVANCE grant agreement. The EMA ADVANCE governance workshop revealed that, in principle, there are no legal barriers to developing a draft contract template, seeking agreement may take time across all stakeholders. However, having such a template would increase the readiness to rapidly initiate studies.

Data sharing with partners outside of the consortium (for validation/replication purposes or other queries) has not yet been tested as not requested so far, but sharing of aggregated data through the current remote research environment is not expected to be problematic with the current infrastructure should this be needed.

All data access providers were able to share data on the basis of the ADVANCE CoC for fingerprinting. DAPs also could provide data to studies using the ADVANCE CoC except for one data source, the regional health agency in Tuscany. The issue was not the ADVANCE CoC or the collaboration with private entities but the extra condition applied by the Agency that a Principal Investigator (PI) (who has overruling responsibilities) of a specific protocol may not be affiliated to a vaccine manufacturer. This restriction applied to one pillar (benefit) but not to coverage and risk ones.

Based on a self-assessment against the Practice Guidance on Quality management, most quality elements were planned and addressed (see annex 2). A formal audit, which is considered to be the ultimate test for adhering to quality matrices, has not (yet) be performed.

Recommendations

- We recommend that the ADVANCE CoC is followed for all collaborative public-private vaccine studies
- We recommend that "conflict of management principles" are written down by workpackage 1 (Best practice guidance) for the post-ADVANCE phase, where real interests may be at stake. Based on the POC experience, a combination of multiple stakeholders in the study team had clear added value. A balanced reflection of different interests (e.g. academic, financial, public health, regulatory) in a study team may avoid undue influence by any specific stakeholder.

- We recommend that public webinars on processes and tools related to the Code of Conduct be organised to increase familiarity with and trust into the system
- A formal quality audit of the implemented ADVANCE procedures to generate evidence might further enhance trust in and acceptability of the system, and guide the need for further changes. An audit based on recommended implementations should be conducted.

2.3 Testing whether data sources are 'fit for purpose' to participate in vaccine coverage, benefits and risk studies

Background/context

The ADVANCE consortium includes full partners and associate partners in 14 countries (UK, Netherlands, Denmark, Sweden, Finland, Norway, Belgium, France, Spain, Switzerland, Italy, Greece, Hungary, Estonia.) A total of 19 databases in 8 countries were accessible for ADVANCE through (associate) partners within the consortium. We applied a distributed network model, where data access providers converted a priori specified select individual level data into common input files, locally.

The main inherent difficulty with collaborative distributed database studies across EU countries has been the heterogeneity of disease coding systems (i.e. ICD-9, ICD-10, READ, ICPC with national adaptations), differences in health care delivery and differences in languages. In order to be interoperable and use the same data transformation code, semantic harmonization and benchmarking is needed.

In order to assess whether the databases were 'fit for purpose' and whether they can be used in vaccine studies, we developed an *ADVANCE eligibility assessment workflow and documentation (details in D5.6)* that included the following components:

1) Information on the meta-data (descriptive characterization i.e. types of database, coding schemes, type data captured) of the database. This was collected through the ADVANCE International Research Readiness Instrument (AIRR survey)[7]

2) Detailed data characterization based on output of standardized analytical code (fingerprinting) of:

- **population** (size over calendar time, distribution of persons by age/sex (and subsequent comparison against national population, availability of birth dates or types of rounding, aggregated entrance and exit dates and follow-up time;
- **vaccinations**: number of vaccines by antigen and age at administration, number of doses, concordance of dose information with the expected one, and estimates of coverage;
- **events**: recorded disease code counts and distributions and associated incidence rates by age, sex and calendar year;

3) National benchmark data on population distribution, vaccine coverage and disease rates for comparison with the fingerprint output from the database.

[7] http://www.advance-vaccines.eu/app/archivos/publicacion/26/ADVANCE_D%203%205%20White%20Paper%20final%20v1.pdf

Based on these three components, it was decided between study teams and the data access provider whether the data source is 'fit for the study (purpose)' (see figure 1)

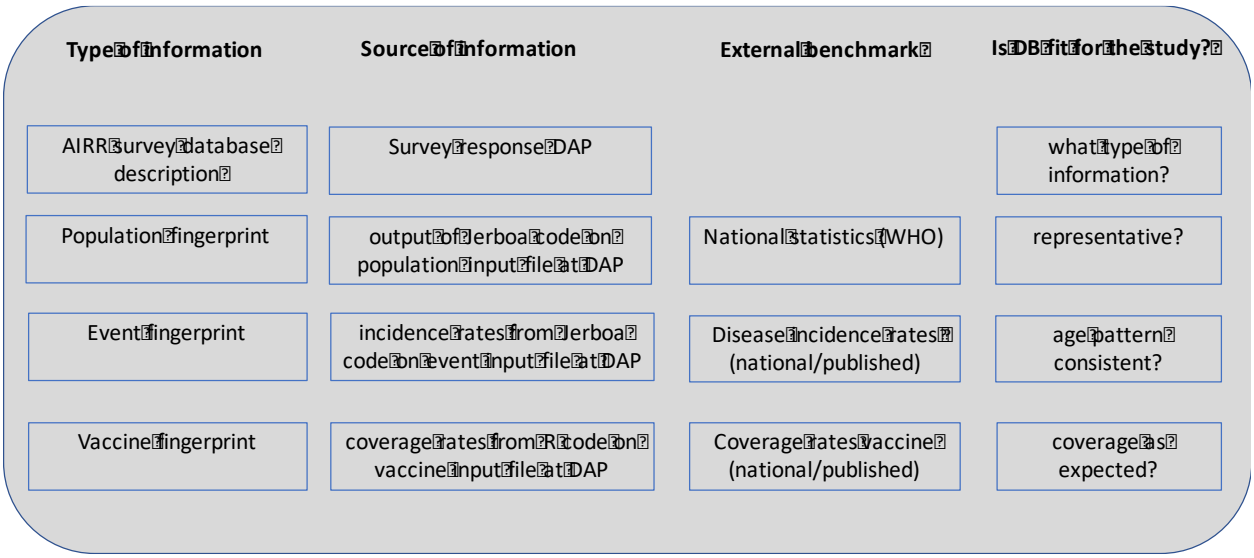


Figure 1: Process for assessing whether the data source is 'fit for purpose'

The AIRR survey was conducted during 2013-2015 with data are available in the ADVANCE data catalogue hosted by EMIF.[8] A first round of fingerprinting of the data content was conducted in 2015-2016 and focused on the events and vaccines of interest for the pertussis proof of concept study. This focused on 10 events and one vaccine (see annex 3). A second round of fingerprinting of database content was conducted in 2017 to improve readiness for potential future studies. In this round, 15 additional events commonly implicated in the vaccine studies were semantically harmonized. In addition, three vaccines (Measles, HPV, Influenza) were included (see annex 3).

Standardized analytical code for transformation of the common input files into the 'fit for purpose' output was: Jerboa (JAVA based open source tool built by EMC, developed and utilized before ADVANCE)[9] and a dedicated ADVANCE R based vaccine fingerprint code.
Information technology tool that was developed and used for semantic harmonization of diagnosis codes was the ADVANCE Codemapper (an open source web application assisting in mapping clinical terms across

[8] https://emif-catalogue.eu/?ref=index
[9] Coloma PM, Schuemie MJ, Trifirò G, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Corrao G, Pedersen L, van der Lei J, Sturkenboom M; EU-ADR Consortium. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. Pharmacoepidemiol Drug Saf. 2011 Jan;20(1):1-11.

different disease terminologies)[10], To address the challenge of coding vaccines we also created an ADVANCE vaccine ontology (VaccO), which is a prototype tool that integrates existing public vaccine ontologies and the Article 57 product database from the EMA. Details about VaccO are described in D5.5.[11]

Learning

Database characterization overall

A total of 19 databases participated (fully or partially) in the fit for purpose assessment, with provision of some meta-data. The type of data captured varied from disease surveillance in Belgium and the Netherlands, to general practice databases, national/regional record linkage data sources and cohort data linked to national registries. The underlying population, type of outcomes captured plus their coding and the availability of vaccines captured differ by provenance of the data sources.

Population fingerprinting

The populations in the different data sources varied from being highly dynamic over follow-up time (GP databases where people move in and out) to very stable (national/regional record linkage databases), but were generalizable to their national reference with regards to age and sex distributions. Several databases rounded the date of birth for privacy reasons.

Event fingerprinting

Event incidence rates in the data sources differed depending on the provenance of the outcome data (i.e. from inpatient and/or outpatient diagnoses and/or laboratory data (e.g. pertussis)). We observed that, as expected, events that are more frequently diagnosed in a primary care setting (i.e. less severe or with more insidious onset such as crying, somnolence) would be less frequently found/have lower incidences in the inpatient setting and the reverse would be true for those events that are more typically diagnosed in the inpatient setting (i.e. more severe such as febrile convulsions).

Vaccine fingerprinting

The fingerprinting exercise of pertussis (DTP), measles (MMR), HPV and influenza vaccination coverage demonstrated that several, but not all, data sources were able to provide vaccination coverage estimates for those vaccines. Most corresponded with national reference data and /or WHO estimates.

Benchmark data

We learned that national benchmarks (reference surveillance data) on events could be found for most of the vaccine preventable diseases, mortality, and for some of the rare and serious events (e.g. VAESCO and PRISM characterization). No EU population-based published incidence rates in countries of interest were

[10] Becker BFH, Avillach P, Romio S, van Mulligen EM, Weibel D, Sturkenboom MCJM, Kors JA; ADVANCE consortium. CodeMapper: semiautomatic coding of case definitions. A contribution from the ADVANCE project. Pharmacoepidemiol Drug Saf. 2017 Aug;26(8):998-1005.

[11] Becker B, Sturkenboom M, Kors J, Martin E, Roberto G. D5.5 Ontology for the Integration and Extraction of Vaccine-related Information in Europe: a proof of concept. http://www.advance-vaccines.eu/app/archivos/publicacion/54/ADVANCE%20D5.5%20Vaccine%20Ontology_v1_clean_noappendix.pdf

found for 'mild reactogenic events': persistent crying, fever, injection site reactions & somnolence, which are relevant in clinical trials but often not a reason to seek care in routine practice and/or symptoms that may not always be recorded. In the absence of published national benchmarks, it was very helpful to compare age specific incidence rates of outcomes between databases within and between countries. Using the provenance of the disease codes (hospital, GP, laboratory, proxy drug) was helpful in further estimating the validity of the outcomes (see below) and quantifying the impact of heterogeneity in the underlying data sources. Extraction of events was relatively straightforward for most events.
Coverage estimates were benchmarked against national surveyor WHO as was the age and gender of the populations against national census data.

Fit for purpose assessment of data sources for proof of concept studies
All 7 databases that completed the first round of events and vaccine fingerprinting in 2016 were eligible to participate in the pertussis POC. The other 12 databases were not eligible at that time because: 1) they did not include pertussis vaccine information (Dutch and Belgian databases, Finnish HPV trial cohorts, Italian Arianna database) or; 2) had access to the required data but did not have the manpower/resources to participate in the first proof of concept studies (FISABIO in Spain, Karolinska in Sweden, Cremona local Italian health agency).

Tools and standardized codes
Semantic harmonization across disease codes was well managed by the ADVANCE Codemapper, or manually, based on that harmonization, if the database used a code system specific adaptation not included in the ADVANCE Codemapper.
The Jerboa tool and output worked well for event characterization and did not create issues for DAPs, most of whom had used this tool in prior EU network projects and could be learned relatively quickly by those DAPs that did not have prior experience. However, we learned that the parameter settings in Jerboa to the study population were not always well known by the database teams, this delayed the understanding and checking process of the aggregated results. So, we recommend full transparency and a priori agreement of parameter settings with the partners involved in the extraction and interpretation of the data. Furthermore, using the provenance of the recorded disease information (hospital, GP, laboratory, proxy drug) was helpful in further estimating the validity of the outcomes (see below) and quantifying the impact of heterogeneity in the underlying data sources.
The R code and processes that were created for ADVANCE were very useful to assess whether data sources can be used for specific vaccine studies, quite some time was spent to define the proper settings for each vaccine and to debug the code, since little R expertise was available and test sets were too limited.

DAP feedback
Data access providers were engaged and enthusiastic to learn how/whether their data could be used in vaccine studies (5 of the 9 that conducted vaccine fingerprint had previously worked on distributed networks for drug safety studies but not vaccine safety studies (BIFAP, SIDIAP, PEDIANET, ARS, ASLCR)). Specific to the vaccination data, appreciable local effort was required to format, clean and prepare these data into the ADVANCE common vaccination input file. Particularly the classification of the antigen included in the pertussis vaccine by acellular- or whole-, required by the first POC proposed, was time

consuming for some databases. Identifying specific vaccines in these databases proved to be a challenge as there was no standardized vaccine dictionary and vaccines are often coded using local codes/languages, and/or they had no recorded dosing schedule or brand/type name. Supporting ontologies/tools such as VaccO, which just became available, will be highly valuable in the future. The ADVANCE R-based vaccine fingerprint allowed for verification of vaccine doses, distributions and coverage estimations using different methods, but DAPs had to familiarize themselves with the settings and participate in the decision of the type of parameters calculated. The discussions between DAP and operations team to evaluate fingerprint data were appreciated and helped in the identification of issues that could be improved, especially regarding the vaccine fingerprint output.

Recommendations:

- The ADVANCE workflows and tools for the characterization and fit for purpose assessment of data sources should be used for future studies to further enhance efficiency and quality.
- The tested procedures should be converted into Standard Operating Procedures.
- The vaccine fingerprint R code might be further optimized to improve user friendliness and applicability to different vaccines.
- Parameter settings in standardized code (Jerboa, R) should be transparent and discussed with DAPs a priori.
- In order to accelerate the process, the exhaustive participation on the methodological design and programming parameter (epidemiological and statistical) is required among the epidemiologist from DAP. In a post-ADVANCE era, DAPs should remain fully involved.
- Results of fingerprinting (beyond AIRR survey meta-data) could be made available in a publicly available data catalogue, preferably hosted by ADVANCE, EMA and/or ECDC for enhanced transparency.
- The workflows and tools might be offered to Daps that were not included in ADVANCE but may have readily available population-based electronic data sources, this would greatly enhance the scale (e.g. Norway, France, Portugal, Ireland, Hungary, Germany, Scotland, Estonia).
- VaccO tools should be further developed to assist local vaccine coding and characterization as well as adding vaccine brand specific information or other information providing information about all levels of vaccine identification (i.e. antigen, components, etc.).
- Further data readiness to generate quality evidence rapidly might be created by
    a. Mapping and fingerprinting all vaccines and additional events of interest
    b. A full transformation into a deep common data model such as OMOP (Observational Medical Outcomes Partnership)[12] or SENTINEL common data model[13], or an adaptation thereof.

[12] https://www.ohdsi.org/data-standardization/the-common-data-model/
[13] https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model/sentinel-common-data-model

Mapping the whole local databases to a common model requires an effort which may prove to be disproportionate to the objective.

- A flexible approach like the one adopted by ADVANCE, where the local data management is the responsibility of the DAPs, may prove to be cost-effective and ultimately provide evidence at higher quality and in a more timely manner, especially if supported by standards and tools specifically developed for European multi-database, multi-national studies.[14]

2.4 Testing the methods to generate data on vaccination coverage, benefits and risks

Background/context

Using simulated data, ADVANCE partners developed several methods relevant to the conduct of vaccine studies using data from multiple countries. Details can be found in D4.4[15]. We tested several of these methods using real world data in the four ADVANCE proofs of concept studies. Methods address outcome measurement, coverage estimations and integration of benefits and risks of vaccines.

Outcome validation is important for scientific credibility but sometimes complicated and time and resource demanding. ADVANCE methods developers demonstrated through analysis of different components of case finding algorithms (i.e. component analysis) that, in theory, chart validation may not be necessary if case validity estimates can be obtained from other sources such as surveillance data, since there is a strong interrelation between different validity estimates. Details can be found in D4.4[16]. This was tested for pertussis.

Outcome misclassification: Electronic data sources have become available for different timeframes in different countries/sites (e.g. SIDIAP data available from 2005 whereas BIFAP data available from 2001). This may result in missing out on outcome data for certain time periods of a study (e.g. data for a period during which whole cell pertussis vaccine was still on the market). Since other databases may have information for the specific time periods tested, the best approach for imputing the missing data was utilizing data from the available databases in the proof of concept safety study.

Coverage

The literature review on coverage conducted by ADVANCE demonstrated that there are no clear standards for measuring vaccine coverage. Electronic healthcare databases could be a good source to monitor coverage, but their key limitation is that the populations covered are very dynamic since individuals are not necessarily registered from birth onwards and may exit the database well before their death. Methods

[14] Gini R, Schuemie M, Brown J, Ryan P. Data Extraction And Management In Networks Of Observational Health Care Databases For Scientific Research: A Comparison Among EU-ADR, OMOP, Mini-Sentinel And MATRICE Strategies. eGEMs 2016; 4.

[15] http://www.advance-vaccines.eu/app/archivos/publicacion/19/ADVANCE_WP4_D4.4_submitted_.pdf

[16] http://www.advance-vaccines.eu/app/archivos/publicacion/19/ADVANCE_WP4_D4.4_submitted_.pdf

to adjust for missing follow-up were created and tested in the ADVANCE proof of concept studies and in the fingerprinting.[13]

Benefit-risk integration
Prior work on methodologies for benefit-risk (B/R) assessments has primarily focused on drugs and devices. At the time of a prophylactic vaccine launch, the benefits clearly outweigh the risks. However, as vaccines' development evolves, quantitative B/R assessment is increasingly becoming recognized as important. In ADVANCE, we explored various B/R assessment methodologies for use with vaccines. A multi-criteria decision analysis (MCDA) was applied in the proof of concept study.

Near real time monitoring
There are several, typically not integrated, facets to post-marketing vaccine surveillance; the surveillance of vaccination coverage, vaccine safety, and vaccine effectiveness and impact. The feasibility and added value of using an interactive visualization dashboard as a potential methodology for near real-time B/R monitoring of pre-specified vaccine health benefits and risks was demonstrated in ADVANCE using simulated electronic healthcare data mimicking the introduction of rotavirus vaccination in the UK[17]. We explored whether near real time monitoring was feasible in the ADVANCE data sources by conducting a detailed survey on identifying relevant time delays that may impede such monitoring.

Learning

We learned that the different methods to deal with electronic health care data and their heterogeneity were implementable and scientifically promising. Estimating the quantitative impact on the incidence rates of applying different case finding algorithms provides transparency and creates an opportunity for adjustment of misclassification, especially when chart review is not possible.

Missing out on outcome data (for estimation of incidence rates) due to the absence of some time period covered or absence of measured outcomes could be handled by imputation of the data using other data sources. Rate or relative risk data could be used from another country, demonstrating the value of the multi-country collaboration.

Real world data showed that estimation of vaccination coverage using highly dynamic populations, especially in general practice databases, is significantly affected by incomplete follow-up over a person's lifetime. This was especially clear in the vaccine fingerprinting, which did not apply in/exclusion criteria. The inverse probability weighting and cumulative density function methods worked well for estimating vaccine coverage under these circumstances, and when available, yielded results comparable to national estimates. We also found that these methods are sensitive to the study population characteristics, in particular age ranges, recommended vaccination schedules, and inclusion and exclusion criteria.

[17] Bollaerts K, Weil J et al. Report on tested methods for vaccine coverage, benefits risks and benefit/risk http://www.advance-vaccines.eu/app/archivos/publicacion/19/ADVANCE_WP4_D4.4_submitted_.pdf

The Multi Criteria Decision Analysis (MCDA) method was implemented, and it generated useful results. Generation of non-linear weights for preferences that are used to compare events, using swing weighting, was difficult for the experts to utilize/operationalize. Experts preferred more linear approaches to create weights. Details are provided in the D5.6 (POC-1 report)

The ADVANCE consortium expressed during the General Assembly Meeting in Barcelona in May 2017, when initial results were presented, that an integrated quantitative B/R analysis may not always be necessary or desirable since examining benefits and risks separately (such as in the dashboard) might be sufficient.

The survey showed that near real time monitoring is feasible across several databases.

Recommendations

- Components analysis[13] and quantitative analysis of different case finding algorithms in electronic health data should be systematically explored as an alternative to chart validation of outcomes;
- The diversity of data sources across Europe, which may cover outcomes differently, should be further explored;
- The variability of results between the various methods used to estimate coverage while adjusting for incomplete follow-up and its dependency on database as well vaccines specific characteristics mandates further research into the robustness of these methods for coverage estimations for each type of vaccine;
- Further work is needed how decision makers like to see quantitative assessments of the benefits and risks of vaccines and near real time monitoring of benefits and risks.

2.5 Testing the platform and work processes to generate evidence

Background/context

Model for collaboration and data sharing

ADVANCE implemented a distributed network model to generate and compile evidence on vaccine coverage, benefits and risks using multiple types of healthcare databases taking advantage of the underlying heterogeneity. This type of distributed network model also creates resource and accessibility efficiencies, for the database providers who may ultimately participate, through the use of a common research protocol and a common data model.

In such a model, the original healthcare data reside and are linked (if needed) at the physical premises of the data access provider, and common ADVANCE protocols and analytics are used to generate and pool

evidence across the data access providers[18]. Data are locally stored, transformed into common data input files, and then aggregated by data transformation code. Results are shared on the central remote research environment, and can be securely accessed. The primary drivers for the selection of such a model were data privacy, quality, and efficiency considerations (i.e. the creation of the study-specific common data model is responsibility of the data access provider, while the embedding of the data analysis in a procedure is the responsibility of the central operations team).

Workflows

The generation of evidence from collaborative distributed network studies requires many processes that were described in D5.3[19]. The numerous steps that need to be undertaken in the execution of protocol-based studies across different databases were translated by the POC evaluation group into a process map (See annex 4) and conceptualized as a **matrix of different steps with many** possible paths depending on the specific needs of each study.

Learnings

A full description of the processes is available in D5.6[20] (report of the proof of concept studies) on the ADVANCE website. We highlight key learnings for the different steps.

We learned that collaboration and engagement across different stakeholders were integral at each of the key steps: study scoping (i.e. selection of the research question)/outline, selection of study teams, protocol writing, analysis and reporting. To be prepared for the future, we used the protocol templates and methods standards for post-authorization safety studies (PASS) recommended by the European Medicines Agency (EMA), and the POC protocols were subsequently registered in the EU PAS Register hosted by ENCePP.

Based on results from the privacy and ethics survey, we learned that none of the participating database providers experienced any hurdles with regards to data privacy and ethics regarding the conduct of studies or sharing of anonymized data. For example, all were allowed to share results of the analyses with the central remote research environment, even if anonymized records were at the individual level.[21] The new

[18] Sturkenboom M, Weibel D, Bollaerts K, Emborg HD, Myint Htar T, Titievsky L, Bauchau V, Haugh M. D5.6 Report on the ADVANCE proof of concept studies. https://goo.gl/Cenaco

[19] Sturkenboom M, Ferreira G. D5.3 POC Research Plan. http://www.advance-vaccines.eu/app/archivos/publicacion/13/ADVANCE_D53_POC_Research_plan_.pdf

[20] https://synapsemanagers-my.sharepoint.com/personal/nyefimenko_synapse-managers_com/Documents/Forms/All.aspx?slrid=71033b9e%2D7055%2D4000%2Dc895%2D53e5c881cde3&FolderCTID=0x0120006CBF727F56BA614FB76D8A043269FE59&id=%2Fpersonal%2Fnyefimenko%5Fsynapse%2Dmanagers%5Fcom%2FDocuments%2FADVANCE%20D5%2E6%20for%20submission%2FD5%206%5FResultsPOC%2D1%5FV1%2E1%5F20170817%5FFinal%2Epdf&parent=%2Fpersonal%2Fnyefimenko%5Fsynapse%2Dmanagers%5Fcom%2FDocuments%2FADVANCE%20D5%2E6%20for%20submission

[21] http://www.advance-vaccines.eu/app/archivos/publicacion/26/ADVANCE_D%203%205%20White%20Paper%20final%20v1.pdf

General Data Protection Regulation (GDPR) rules entered into force on 24 May 2016, and will fully apply from 25 May 2018. A key impact of these new rules for future vaccine benefit-risk studies include the appointment of a Data Protection Officer in every organization dealing with re-use of sensitive data.

Data management and analysis were directed by a cross-study *analytics team* of 9 persons that was created to permit knowledge sharing and ensure alignment. Data transformation scripts were prepared by the analytics team who also wrote the statistical analysis plan, both in R and in SAS, since some databases do not have SAS licenses. The data transformation was developed on the RRE so that participants could share codes, and these scripts/codes were then sent out to the DAPs to run on their local data. We learned that the creation of the data transformation scripts took a long time. Problems encountered were: 1) limited number of persons able to code in R; 2) difficulty getting access to the RRE for vaccine manufacturers who experienced difficulties installing the token application on their company systems; 3) unfamiliarity with the processes and distributed analytics methods; 4) limited time to dedicate to ADVANCE; 5) insufficient space on the RRE and the time it took to run prototype scripts/codes; and 6) limited time and process for quality control and verification of the code and outputs between SAS and R. 7) Limited test set which resulted in need to debug when code as sent out to DAPs

The common challenges stemmed from several sources: 1) most persons in the consortium had subject matter expertise in vaccines but were unfamiliar with distributed analytical processes, the common data model, and the remote research environment; 2) few had prior experience in conducting or coordinating distributed multi-site studies or benefit-risk analysis. The implications of this included the “unease” in providing scientific input on required areas of the POC, and time delays in decision making that were exacerbated by working in a virtual environment vs. in person and not having had worked together in the past.

Recommendations

- Readiness of vaccine researchers, statisticians and epidemiologists to take specific roles in distributed studies should be improved with dedicated training programs on the processes, templates, tools and quality requirements;
- ‘bootcamps/scrums’ for study teams, analytics team and database team should be organized to enhance speed and consensus finding;
- Dedicated study/project management is needed to keep timeliness;
- A future transient data sharing platform should be secure, and allow for secure but easy access and have large computing facilities, preferably a private cloud service;
- A data privacy impact assessment (DPIA) on central environment should be done to be compliant with GDPR. We recommend that a test DPIA be done (by EMC) on the current RRE to identify the hurdles we may face in May 2018;
- Future studies should test the ability to generate evidence faster to improve acceptability.

2.6 POC Evaluation

A specific study team was created for the evaluation of the POC studies. The study process map that was created was based on a review of all deliverables, as well as a discussion with the study coordination team; this process map could guide the design of future studies (Annex 4). According to ADVANCE stakeholders, ADVANCE workflows have not yet demonstrated significant change in the speed of getting approvals, but the tested processes may be transformed into standard operating procedures (SOP) including a "fast-lane" process for data clearance in case of a crisis. The system is expected to be ready for production regarding data access from the current databases (partners in the ADVANCE consortium). Overall, the ADVANCE solution is deemed scalable, and can be applied to address other B/R questions for other vaccines and risk or disease areas.

Recommendations

- The process map should be used to plan future studies and run parallel tracks;
- ADVANCE should continue testing the platform and optimizing the processes which could be transformed into SOPs;
- System evaluation should be conducted using specific target groups and face-to-face feedback sessions, since in large consortia not all people are aware of all steps.

3 The future and next steps

Through the different pillars comprising the proof of concept we have tested the ADVANCE system, which proved to be feasible for generating evidence on vaccines benefits and risks. In addition, the ADVANCE system was deemed acceptable by different stakeholders.

Main areas for improvement are:

1) Reduce throughput time by parallel process planning, and improving people readiness;
2) Providing clarity and training about the procedures for all stakeholders;
3) Expand standardization of procedures and workflows. For example, this may include but not be limited to the use of the VaccO, further improvement of vaccine fingerprinting, semantic harmonization and fingerprinting of more events of interest;
4) Inclusion of data sources beyond the consortium;
5) Be prepared and test future challenges such as GDPR;
6) Optimize study team members' engagement, commitment, accountability and decision making across large stakeholder group working in a remote working environment;
7) Conduct a formal quality audit to identify areas where quality may be improved;
8) Conduct a real 'safety study' to test the Private Partner Collaboration (PPC).

It is recognized that only selected types of databases were included in the POC studies, which was a function of who answered the call for participation/collaboration in the consortium, and/or available resources/ability to commit to the execution of these studies at the time of the request for participation. Several relevant types of databases that had to be excluded or not fully evaluated at the time of our proof of concept study are insurance and diseases surveillance databases. Based on the processes and the fingerprinting of those source that have been tested to date, the process for evaluating such databases in the future is relatively straightforward and would include the steps below bearing in mind the research question(s) of interest:

1. Obtaining information on the structure of meta-data of these databases.
2. Conducting data characterization/fingerprinting.
3. Comparing the data generated in step 2 to that of relevant country benchmark.

In conclusion, we demonstrated that in the case of vaccine studies:

1. Collaboration across different types of stakeholders is feasible though trust and mutual respect, as well as an understanding of individual objectives and the context within which partners and stakeholders' work. Indeed, ADVANCE provided an environment where it was possible to create and grow this trust while working on system testing; the collaboration did not yet test a real safety issue.
2. ADVANCE has operationalized the ADVANCE Code of Conduct, data sharing principles, ethics and privacy specific to the context and concerns of EU countries that are similarly aligned with the above mentioned models in order to implement and test a prototype of such a network [distributed network model can work in theory but the ability to replicate such a network, and sustain and expand it, ultimately depends on the commitment of countries and DAPs to engage long-term under a pre-specified and agreeable set of conditions (beyond the immediate scope of ADVANCE)];
3. Data characterization work flows are useful and efficient for examining potential data sources and assessing the feasibility for inclusion in a potential study.

The ADVANCE system is ready for testing near real time monitoring, real vaccine B/R questions and production.

Annex 1: Recommendation in the ADVANCE CoC and implementation in the POC

Topic	CoC Recommendations to be applied in all studies ("must")	CoC Other recommendations to be considered for all studies ("should")	Practical Implementation in POC study
1. Scientific integrity	1. All study team members are qualified to fulfil their role		1.Members applied for roles with CV. Qualifications reviewed by decision making organ
	2. All study team members act in accordance with core values of honesty, accuracy and objectivity		2. Compliance statement in application letter for study team
	3. Study team members adhere to IEA Good epidemiological practice and ISPE Good pharmacoepidemiological practices		3. Compliance statement in application letter for study team
2. Scientific independence	4. Study is conducted without undue influence of any financial, commercial, institutional or personal interest in a particular outcome of the study	5. Autonomy of members of study team for making scientific decisions in their organisation is documented	4. Declarations of interest signed and made publicly available by all members 5. Was not documented in PoC
	6. Scientific independence is safeguarded by clear and transparent roles and responsibilities, peer review process, transparency measures and disclosure of all funding sources		6. Full transparency on DoI, roles and comments /edits throughout the study. No study member can decide alone. Tracked change documents and rebuttals on Sharepoint & detailed document histories
3. Transparency	7. Study is registered in a publicly accessible database before the start of data collection or extraction	11. Final study report or summary is uploaded in publicly accessible database where study is registered	7. Studies registered in ENCePP registry 11. Final report in ENCePP registry and on ADVANCE website
	8. Sources of funding are made public at the time of registration, in the protocol and in the presentation of results	12. After study completion, study information is made available from outside the study team in a collaborative approach	8. IMI funding is made public 12. Study information available on ADVANCE public website
	9. Declarations of Interests (DoI) are made available at an early stage of the study, regularly updated and disclosed	13. In case of primary data collection, participants in the study or their representatives may receive main study results and interpretation thereof	9. Declaration of interests for each study team member publicly available on ADVANCE website 13. Not tested only secondary use
	10. All comments received on study protocol and results with impact on the study are documented		10. Detailed track changes and comments plus rebuttals on protocol and report publicly available

4. Conflict of interest	14. Actual or potential conflicts of interest (and perceptions thereof) are addressed at the planning phase of the study. Research contract includes a description of the management of conflicts of interest. All DoI are made publicly available	15. A standard form is used to declare all interest that may lead to conflicts	14. DoI publicly available. No process for dealing with potential conflicts of interests was tested. 15. DoI were collected in a standard form
5. Study protocol	16. A study protocol is drafted as one of the first step in any research projects	18. The process for reaching an agreement on the design options of the study is agreed beforehand	16. study protocol drafted 18. Design choice decided first in study team and subs. By Steering committee
	17. Study protocol is developed by persons with relevant expertise	22. Detailed draft protocol undergoes independent scientific review	17. Study team composition had clear entry criteria (epidemiology PhD, vaccinologists, disease experts, statistician) 22. Draft protocol reviewed by SC and Scientific Advisory Board
	19. Protocol includes a section with ethical considerations involved and information on funding, affiliations, potential conflicts of interest, data protection and incentives to subjects. Protocol is approved by relevant research ethics committee	23. Protocol is registered in publicly accessible database before the start of data collection	19. EMA PASS protocol template used with ethical, privacy & regulatory section 23. All protocols in ENCePP database before start of data collection
	20. Protocol includes description of each party to study design, protocol writing and work programme		20. Detailed document history in Protocol and study team roles/responsibilities
	21. Regulatory obligations and recommendations applicable to the study are described		21.Section on regulatory obligations stated in protocol
	24. Changes to the protocol that may affect the interpretation of the study are identifiable and reported in the study report		24. All amendments listed in report. Report also has limitation section
	25. Key statistical analyses are described		25. Detailed Statistical analysis plan provided separately and public, summary described in report
6. Study report	26. Set of principles are followed for reporting results including documentation of important safety concerns and deviations from protocol or statistical analysis plan, sources affecting data quality, strengths and limitations, and sources of funding	27. STROBE statement and internationally agreed guidelines are consulted when analysing and reporting data	26. Reporting results according to EMA PASS report template, including limitations, data quality and funding 27. STROBE and RECORD guidelines followed in report
		28. Draft study report undergoes independent scientific review	29. Study report evaluated by SAB, SC and in house
		29. Study report or summary of the results is included in the publicly accessible database	29. Study report is available online (ADVANCE website) and summary in ENCePP study registry

7. Publication	30. All study results are made publicly available. Authorship of publications follows the rules of ICMJE	32. Preliminary or partial results of discontinued study are reported and identified as such	30. Report publicly available & in line with ICMJE 32. Monitoring of coverage, risks and benefits not done, but part of POC1.2 (was described)
	31. Research contract allows the principal investigators and relevant study team members to publish study results independently from the funding or data source. The study requester/funder may provide comments		31. by IMI grant agreement & consortium processe, publication rights.
	33. Procedures are in place to rapidly inform regulatory and public health authorities of study results, independently from submission of a manuscript		33. Communication procedures in place and less relevant because all stakeholders are involved
8. Subject privacy	34. Privacy of study subjects in relation to personal data is core principle of any medical research		34. Study in compliance with relevant EU directives
	35. In case where personal data are collected, the applicable legislation is followed		34. Study in compliance with relevant EU directives and national implementation thereof. Detailed data collected through P& E tool
9. Sharing of study data beyond study	38. Sharing of study data is based on a written request specifying the ground of the request. The study team verifies the compliance of the request with the data protection legislation	36. There is an open and collaborative approach to sharing study data with persons from outside the study team	36. Study results shared publicly. Data can be accessed through RRE after signature of confidentiality & security documents 38. Has not occurred but processes in place
	39. Requests for data sharing are justified based on public health interest	37. Data are shared only after the study report is finalized	37. Has not occurred 39. Has not occurred
	41. Analyses performed with shared data follow the provisions of the ADVANCE Code of Conduct	40. Study team or delegated committee takes the decision to share study data	40. Has not occurred 41. Has not occurred
10. Research contract	42. The research contract does not lead investigators, directly or indirectly, to act against the principles of the Helsinki Declaration or applicable legal or regulatory obligations	44. Unique multiparty contract is preferred in cases where several parties interact	42/44. Could not be tested due to IMI context
	43. Clarity and transparency are key elements of the research contract	45. The research contracts indicates that the study will follow the ADVANCE Code of Conduct and provides core information	43/45. Could not be tested due to IMI context

Annex 2: POC Implementation of quality framework as queried by WP1

Item	Criteria	POC Implementation
Study protocol development and protocol deviations	Documented expert review of protocol	All protocols approval by POC CT, SC and SAB
	Recommended or standard template for protocol	EMA template
Reporting and communication of study results	Documented expert review of study report	By POC CT , SC and SAB
	Recommended or standard template for study report	EMA template
Qualification of external vendors/service providers/data providers/analysts (if applicable)	Documented process of vendor qualification	Study teams put together based on pre-defined requirements, all CVs provided
	Defined criteria for assessment of vendors	Standardized Data 'fit for purpose' assessment and documentation
Ensure independent research in case of unconditional grant	Written documentation of scientific independence	Declarations of interest collected from all study team members and publicly available
Protection of data on premises, servers and individual work stations	Use of security logs	local responsibility RRE is logged
	Designated and controlled areas for data storage	Data controller responsibility local sites, and central
	Access to data only for authorized personnel	Centrally organized on RRE, locally data controller
	Log on with multi-character passwords	RRE organized (token)
	Data storage after study end for minimum of five years	Local and central storage
	Data storage index present for audit and inspection purposes	Storage and archiving rules on RRE by origin and date
Protection of identifiable and confidential data	Replacing of overt personal identifiers by clear identifiers, keeping the mapping key separate from the pseudo-anonymized data	Locally according to national rules and EC directive. Centrally only anonymized data
	All researchers must sign a confidentiality agreement	To access RRE researchers have signed CA

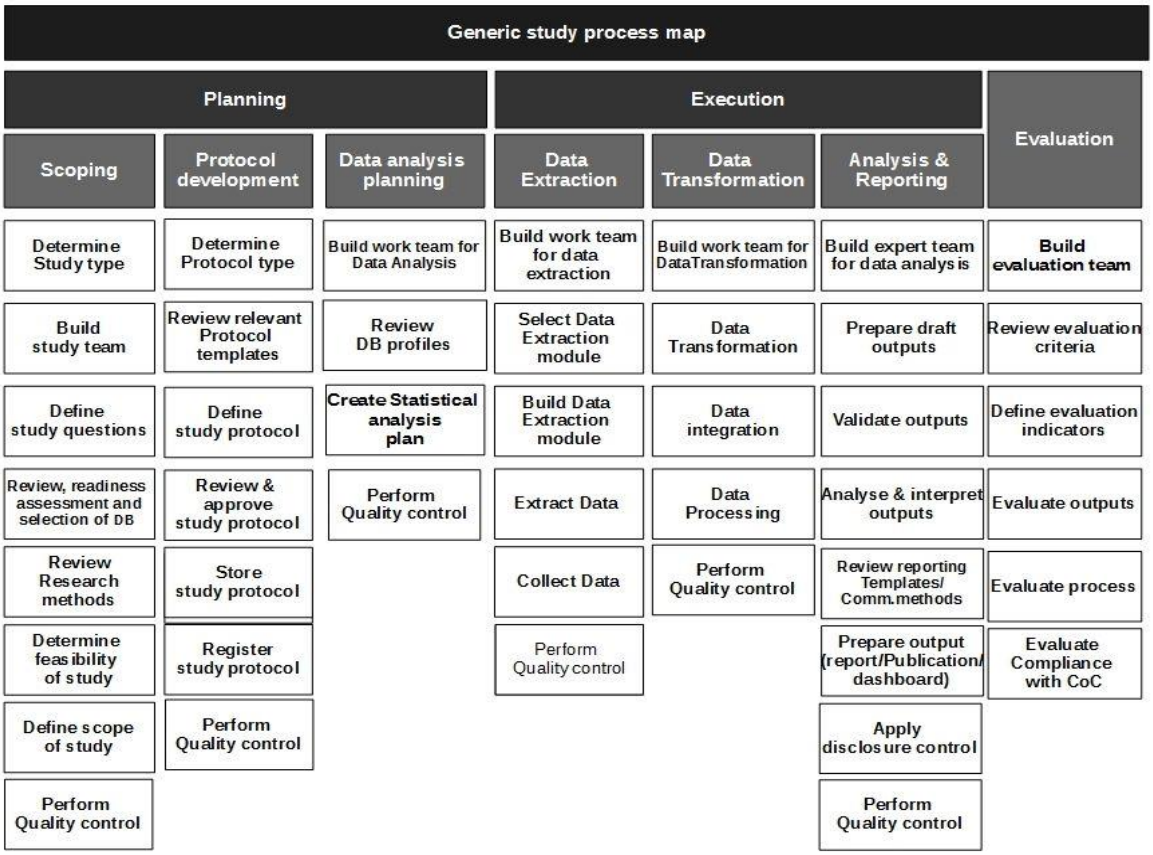
Back-up	At least one (but preferably multiple) back-up(s) in different location(s)	RRE back up on mirror server
Process of electronic data transfer	Only sent data from one place to another by secure methods (encrypted)	From local to RRE through ftp line
Data processing (and statistical programming if applicable)	System-generated audit trails in place	Local and central trails available (not necessarily system generated)
	Application/execution of standard consistency and accuracy checks	Fit for purpose assessment and workflow
	Availability of Statistical Analysis Plan (SAP) prior to data analysis	SAP available for each pillar and attached to report
	Use of validated statistical software	SAS & R
	Annotated study programming (programming giving explanatory notes for each step)	Conducted
	Standard process for statistical programming control (i.e. review or double-programming)	Double programming
	Archival of SAP and statistical programs	SAP is archived on Sharepoint, programs on RRE
Organization and responsibilities for data privacy	Formally documented that data is legally obtained	Ethics/Scientific committee review documentation on Sharepoint
	Presence of (an) allocated person(s) responsible for data privacy	Data privacy person is local responsibility
Ethical review board	Process for obtaining Ethics Committee approval on an appropriate level	Local processes have been identified a followed
Informed consent	Obtain ethics approval of informed consent or waiver of informed consent	Local processes identified by WP 3, in line with local practice
Ensure sufficient qualification of study personnel	Ensure a principal investigator is qualified and appointed	PI needed to be PhD, epidemiologist and have proven experience to guide multisite studies
	Have written organisational charts and personnel tasks in place	Roles and responsibilities defined in D5.3
Training system	Initial and continued training of personnel	Continued Capacity building within ADVANCE by WP5 leads and participants
	Continuous documentation of training status and certification	CV of investigators are collected
Commitment	Allocation of resources and qualified personnel prior to study start	Public call for people
Existence of quality cycle	Continuous cycle of planning, adherence, control & assurance, improvements of all processes in place	POC EC

Written policies and procedures for main processes/activities and systems	Quality plans on procedures	WP 1 working group
	Quality manuals for procedures	POC Research plan (D5.3).
	Periodic review and update of procedures	Continuing development, no adjustment done to D5.3
	Record management policy	To be developed
	Urgency processes/escalation policies	Not available
Document control and document management	Review and approval process of documents	As per ADVANCE project handbook and communication guidance
	Trackability of records (version control processes / timestamps)	Version control used but not standardized
	Controlled document management system	No
Audit and inspection preparedness	Inspection plan available	no
	Periodic internal audit	POC EC and SC?
	Documentation of audit reports and results	Not done
	Periodic check of study facilities and equipment	Not done
	Deviation management; corrective and preventive actions and follow-up	Not done
Ensure adherence to procedures/compliance management	Written compliance management process	Not available

Annex 3: Aggregated fingerprint data provided to the remote resource environment from different data sources

Overview of available fingerprint data on the RRE (green means available on RRE and quality verified, red means not available in database, white means not extracted (resource issues), orange: data available but still under review for consistency)

Database	General practice databases							Claims & hospitalizations linkage					In-outpatient linkage		Disease surveillance				
Country	NL	IT		ES		UK		SE	DK		IT		ES	FI	BE				NL
DB code	IP	PE	CA	BI	SI	TH	RC	KI	AU	SS	AS	AR	FI	CO	SG	PE	NR	SL	OS
Period covered by data		2003-2016		2003-2015	2005-2016	1985-2016	1989-2016	1998-2002	1995-2016		2001-2016	2002-2016		2006-2015	2008-2017				
Pop. (Million)	1.1	0.075		6.9	6.8	11.6	1.8	9.4	7.4		0.4	4.5		0.065	0.1				16
Vaccines																			
Pertussis																			
HPV																			
Measles																			
Influenza																			
Events																			
ADEM																			
BP																			
Convulsion																			
Feb. conv.																			
Gen. conv.																			
Death																			
Fever																			
GBS																			
HHE																			
ISR																			
ITP-broad																			
IT_narrow																			
Kawasaki																			
Narcolepsy																			
Optic neu.																			
P. crying																			
Pneumonia																			
SLE																			
Somnolence																			
Transv. M.																			
Pertussis																			
Measles																			
Meningitis																			
Pneu.Menin																			
Mumps																			
ILI																			
Possibility NRTM	no	yes		no lag time > 1 year	yes	not with EMC license	yes		yes		yes	yes		no					

IP=IPCI database; IT-PE=PEDIANET database; CA=Caserta ARIANNA database; BI=BIFAP database; SI=SIDIAP database; TH=THIN database; RC=RCGP database; KI=National Swedish data provide by Karolinska Institute; AU=Aarhus databases; SS=SSI national DK databases; AS=ASLCR database; AR=ARS (regional data Tuscany); FI=FISABIO; CO=Cohort; BE-SG= ILI surveillance by Belgium Sentinel General Practices, BE-Pe= PediSurv is a Belgium children infectious diseases network of GPs and paediatricians), BE-NR=Belgium National Reference Center; BE-SL= Belgium Sentinel Laboratories; BE-NS=national surveillance; OS=Osiris. NRTM: near real time monitoring of vaccine benefits and risks (ie biweekly/monthly)

Annex 4 ADVANCE Process map