D4.8 Data Management Plan

777363 – DRIVE

Development of Robust and Innovative Vaccine Effectiveness

Lead contributor	Tom De Smedt (3 – P95)
	tom.desmedt@p-95.com
Other contributors	Margarita Riera (3 – P95) Kaat Bollaerts (3 – P95)
Reviewers	Mendel Haag (14 – SEQIRUS) Gael Dos Santos (15 – GSK Bio) Alexandre Descamps (18 – INSERM)

Due date	31 DEC 2019
Delivery date	19 DEC 2019
Deliverable type	R
Dissemination level	PU

Description of Work	**Version**	**Date**
	V2.2	18 DEC 2019

Document History

Version	Date	Description
V1.0	22 DEC 2017	Final Version for D4.2
V1.1	17 JUN 2019	Updated Version for communication with QCAC after Season 2018-2019 conclusion
V2.0	06 DEC 2019	Updated Version for review for D4.8
V2.1	18 DEC 2019	Updated Version after review round 1 for D4.8

The Data Management Plan is intended to be updated regularly and is considered a living document. As reflected in the document history above, the document is both D4.2 (DMP in the state as of end DEC 2017) and D4.8 (DMP in the state as of end DEC 2019).

Table of contents

1 Introduction and aim

The Data Management Plan (DMP) provides a description of the data management that will be applied in the DRIVE project including:

- A description of the data repositories, who is able to access the data, and who owns the data.
- The main DMP elements for each of the studies contributing (or sharing data) to DRIVE.
- The time period for which data must be stored.
- The standards for data collection, validation and evaluation.
- The possibilities of and conditions for sharing data.
- The implementation of data protection requirements.

The DMP will be updated over the course of the project whenever significant changes arise, such as (but not limited to):

- Addition of new data
- Changes in consortium policies (e.g. new innovation potential, …)
- Changes in consortium composition and external factors (e.g. consortium members and/or associated partners joining or leaving).

In summary, the DRIVE DMP gives guidance and provides an oversight of general data management, while each study needs to provide specific data management information including, but not limited to, data capture systems, data analysis systems, data protection and data privacy measures, including description of de-identification of data sets and access rules. In cases where the research results are not open access a justification needs to be provided.

2 General principles

The DMP is a working document, that will evolve during the DRIVE project, and will be updated to reflect project progress.
The DMP follows the principles that research data are findable, accessible, interoperable and reusable (FAIR)[1].

The general principles on access rules are defined in the Consortium Agreement (Section 8 Intellectual property – Access rights).

[1] European Commission Horizon2020 programme. Guidelines on FAIR Data Management, v3.0, 2016. (http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

3 Data flow

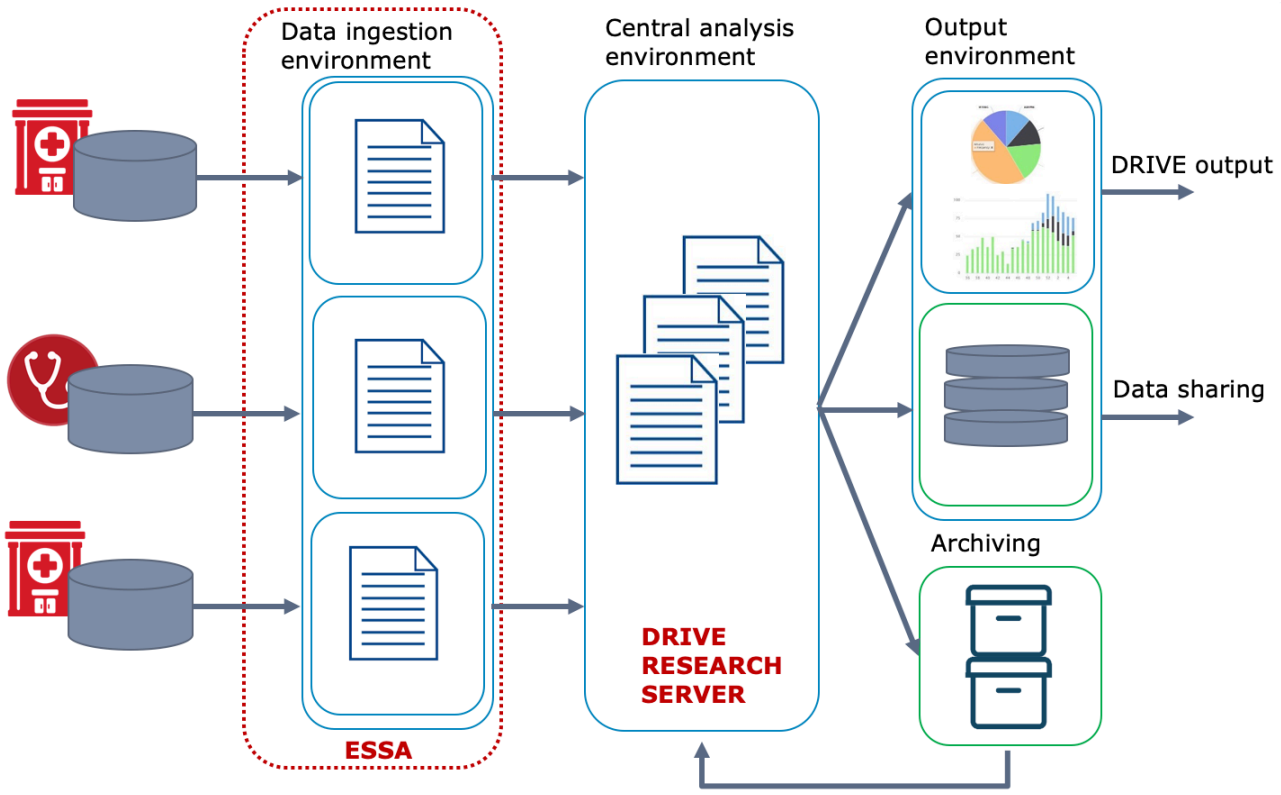


Figure 1 - Data flow from study site to DRIVE Research Server and beyond

The data flow is as follows, after a Statistical Analysis Plan (SAP) for a study is accepted:

1. Each local site data custodian creates a selection of the database as per the SAP and the Minimal Dataset Requirements (MDR)
2. Each local site data custodian uses the Electronic Study Support Application (ESSA) webtool to upload the selected data and perform quality checks and process any findings accordingly with sufficient documentation to ensure reproducibility
3. When the performed quality checks are satisfactory, the data custodian can flag the loaded data for ingestion in the DRIVE Research Server (DRS). No data is stored on the ESSA server after the data custodian leaves the web tool
4. The flagged data is automatically send from the ESSA to the DRS using sFTP into a folder that is solely dedicated to the DRS server admin
5. The DRS server admin checks whether the data is in compliance with the SAP, MDR and privacy regulations
    a. If the check is satisfactory, the server admin releases the uploaded data to the study folder accessible by the data analysts (using Remote Desktop Protocol) and performs a data lock (the “raw” data is only readable by the data analysts and cannot be changed)
    b. If the check is not satisfactory, the server admin reports this to the data custodian responsible for the data
6. The data analysts perform the required data transformations on the data released in the study folder as per the SAP
7. When the data transformations are finalised, the data analysts flag the resulting output files to the server admin for extraction out of the DRS. These output files only

contain highly aggregated summary data such as figures, tables with number of events, or estimates.

8. The server admin checks the resulting output files flagged for extraction for compliance to the SAP
    a. If the check is satisfactory, the resulting files are extracted out of the DRS by the server admin using sFTP
    b. If the check is not satisfactory, the server admin reports this to the data analyst and requests changes to get into compliance with the SAP
9. After the resulting files are extracted out of the DRS, they can be used as the basis for reports, web applications, etc. as per the SAP

Modalities for archiving and data sharing are further discussed in Sections 6.1.5 and 6.10

4 Statistical Analysis

Tabular and graphical descriptive summaries of the site-specific and combined data are obtained. A meta-analytic approach is used to obtain crude and confounder-adjusted influenza vaccine effectiveness (IVE) estimates by any influenza vaccine, vaccine type and vaccine brand against any influenza, influenza by type and by subtype/lineage. The meta-analytic approach is a two-step approach;

- First, all site-specific confounder IVE estimates are obtained using logistic regression (for case-control studies) or Poisson regression (for cohort studies).
- Then, in a second step, the site-specific IVE estimates are pooled across sites using random effects meta-analysis

Both steps are conducted centrally using common analytical tools and scripts. This way, the conduct of the statistical analyses is accelerated and the harmonization of the analyses across sites maximized. The details of the statistical analysis can be found in the SAP provided in Annex 2 and in D4.4 Generic Statistical Analysis Plan. Quality control (QC) of the central analytical tools and scripts are performed through code review by an independent statistician. The findings of the QC of the analytical scripts are summarized in Annex 3.

5 Overview of data managers and access rules

Two data repositories/platforms are used in the DRIVE project:

- One platform, the Electronic Study Support (web)Application (ESSA), which will be used by the data custodians to upload data and perform data quality checks
- One data repository, DRIVE Research Server (DRS) hosted by P95 which will be used to store all study related datasets produced and/or shared within WP7, with limited access.

The contact details for the data management team are described in Table 2.

Table 2. Data management team contact list

Responsibility	Name	E-mail address
Data management	Tom De Smedt	tom.desmedt@p-95.com

compliance / Server admin		
Deputy data management compliance	Roberto Bonauiti	roberto.bonauiti@unifi.it
Deputy data management compliance	Cintia Munoz Quiles	munoz_cin@gva.es

The following generic user roles will be defined:

- Study lead
- Data owner
- Data analyst
- Server administrator

It is possible for a single person to take on multiple generic roles.

- The study lead is responsible for the conduct of the study, ensuring the adherence to the protocol, statistical analysis plan (SAP) and study procedures.

- The data owners are the database custodians and are the responsible at the local data sources that have the necessary data to participate in the study. The local data providers are responsible for extracting the requested data out of their database(s) according to the study protocol (including the MDR) and study-specific DMP and importing the extracted data in the ESSA.

- The data analyst uses the imported data to perform the necessary data transformations following the protocol and pooled analysis SAP in the central analysis environment. The resulting output on aggregated data (tables, figures) is exported from the DRS for further analysis and communication as per protocol and SAP procedures, led by the study lead within WP7.

- The server administrator of the DRIVE Research Server (DRS) is responsible for the following tasks:
  - Set-up and maintenance of the DRS (permissions, security, logs, updates, etc.)
  - Responsible for user management (registration of local data providers and data analysts, two-factor authentication, connection set-up and monitoring)
  - Validate uploaded data from local data providers
  - Transfer files between DRIVE server compartments (ESSA, central analysis and output)
  - Validate files flagged for export by data analysts
  - Perform privacy assessments where necessary with every data transfer step

Access to the DRS will only be granted to DRIVE consortium members, excluding EFPIA partners, and associated partners that are part of WP7, or part of the Quality and Audit Committee (QCAC) as defined in the CA. Access can be requested by completing an intake template document, detailing necessary information about who is requesting access (i.e. a data analyst, etc.) and the purpose (for which study, for audit purposes etc.). After DRIVE steering committee decision to grant access, this document should be sent to the server admin, who will grant the access. The intake document also contains contact details necessary for getting access with two-factor authentication. The intake template document

can be found in Annex 1.

6 Technical specifications

6.1 DRIVE Research Server

A secure repository to store datasets is provided by P95. Within this repository, research sites will be able to upload study specific datasets in order to be able to perform pooled data analysis using the ESSA. This is a highly secure environment and network, with strict rules for data access. The infrastructure is in accordance to the new GDPR guidelines on storing personal identifier data and in a processor-role, as well as storing anonymized data. For more info, also refer to Section 11.

6.1.1 Specifications

The repository consists of:

- Dedicated secure virtual server on redundant cluster
  - Physical connectivity: 100Mbit/s
  - Traffic: 100GB/month
  - Hardware: 4 CPU cores
  - 8 GB RAM Memory
  - 500GB Storage Memory
  - Location: Belgium
- Backup storage platform and necessary licenses
- Installed software and licenses
  - Windows Server 2016
  - MS Office 2016
  - Microsoft SQL Server
  - Remote desktop licenses
  - R
- Monitoring of operating system
  - System update status
  - System log files
  - Access to system, scheduled batch job status
  - Memory
  - Disk space and status
  - Monitoring status
  - Backup status
  - Running services
  - Additional application maintenance on the webserver configuration and SSL certificate status
- Server availability Monitoring & Reporting (SLA)
  - 24/7 monitoring and reporting of system availability and security
  - Weekly server security scan and necessary corrective actions
- Two-factor authentication using DUO Access®
- Additional server set-up documentation
  - Universal Routing Server (URS) document
  - Workflow specifications (functional, non-functional & design specifications)

- Workflow validation

6.1.2 Procedures/tools for data accessibility/security

Details of all users of the repository must be registered with P95. Access will be restricted to a minimum number of people needed for analysis, code review or supervision of the study.

The architecture of the DRIVE Research Server is shown below (Figure 1).

The described IT architecture allows for all types of data (i.e. individual record level data, anonymised record level data and aggregated data) to be used and allows for multiple studies to be carried out at the same time.

The proposed IT architecture is:

- Scalable
- Secure
- Transparent
- Can use all formats of data
- GDPR-compliant

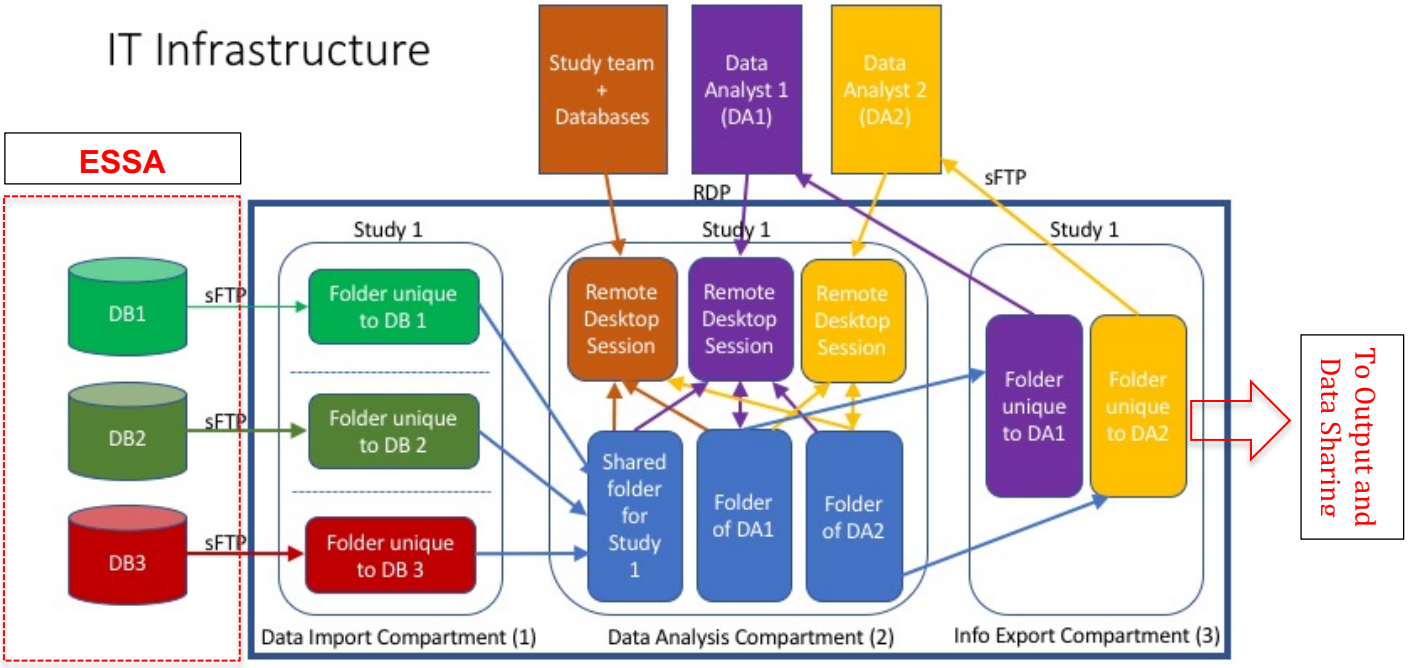


Figure 2 - Architecture of the DRIVE Research Server (DRS)

The general architecture of the DRS has three compartments: the data import compartment, the data analysis compartment and the info export compartment. The DRS is only accessible through the secure file transfer protocol (with upload capability to the data import compartment and download capability out of the data export compartment) and the remote desktop protocol allowing (primarily) the data analyst role to log into the data analysis compartment. The transfer of any data between the different compartments is done solely by the server administrator role, where data privacy assessments should be carried out if deemed necessary. Every interaction on the DRS will be logged, and these logs will be accessible upon request.

6.1.3 Duration of accessibility

Authorized users will have access to the DRS during the course of the DRIVE project, and as long as the user is a member of the project.

6.1.4 Back-up process & Disaster recovery

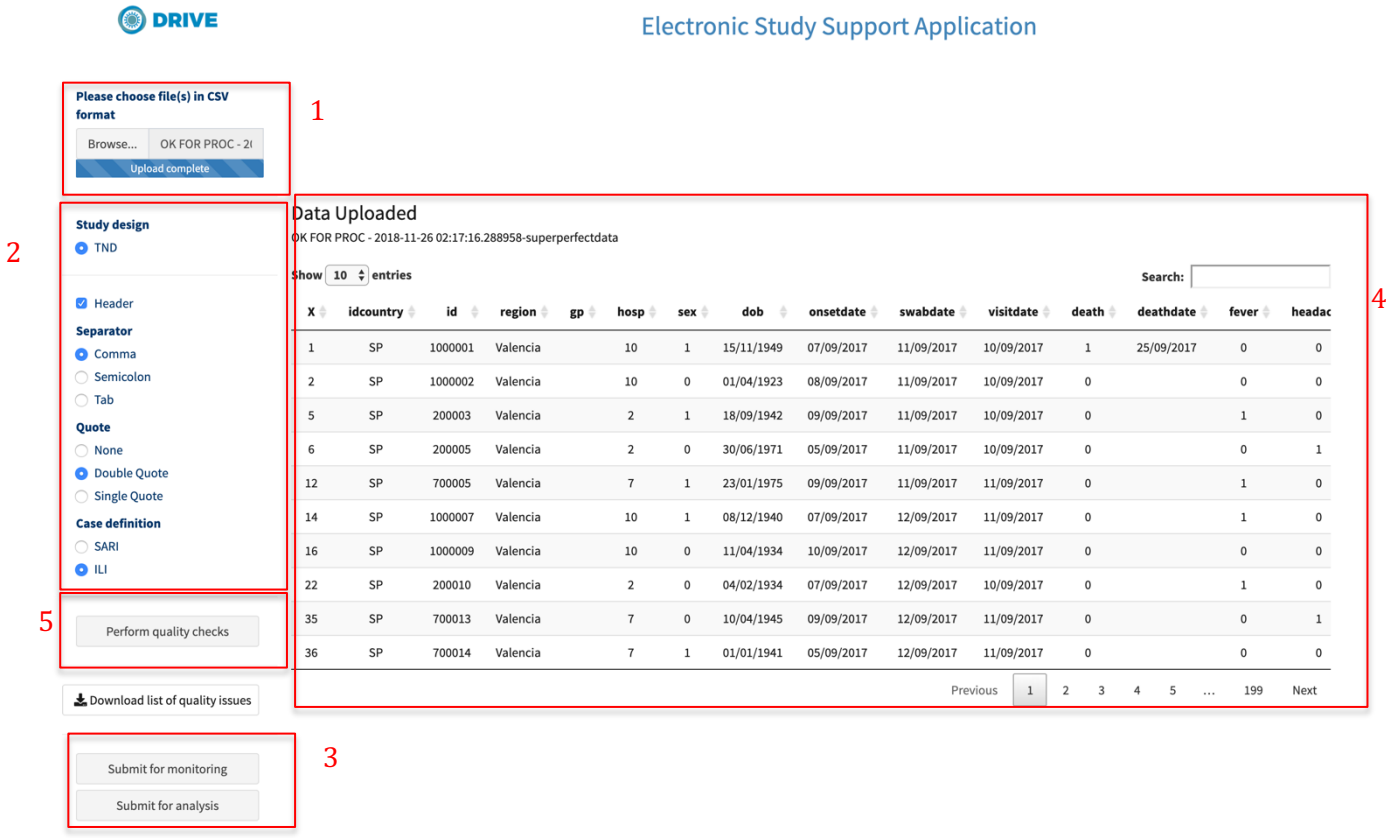
There will be on-site and off-site backups of the server OS, the application code and data:

- Daily system block level encrypted backup
- Retention of 30 daily backups
- Reporting on yearly manual backup restore test of server and on deleted backups on request
- Backup locations:
  - Interxion Datacenter, Wezembeekstraat 2, 1930 Zaventem
  - Uniweb BVBA, 's Herenweg 16, 1860 Meise

6.1.5 Archiving and preservation

All study related data in the DRS produced as the results of the DRIVE project will be kept for 5 years following the completion of the DRIVE project as specified in the Consortium Agreement (Article 18). The details of the physical location of the archive (if different from the location used during the project) after the completion of the project as well as access will be developed in future versions of the DMP.

6.2 Electronic study support (web) application platform (ESSA)



In the ESSA, the data custodian can upload their data using an upload button (1) and by setting the parameters (header yes/no, which separators, etc.) for a correct upload (2) according to the specific format of the uploaded data. Upon upload the uploaded data is shown in a table (4) which the user can use to explore the data. When clicking on the quality check button (5), more quality checks (on missing data, order of dates, etc.) are performed and shown using HTML tables and interactive charts. When the quality checks are deemed satisfactory, the user can flag the uploaded data for analysis (3), meaning it gets automatically send to the DRS using sFTP. Upon log off of the ESSA by the user, the uploaded is deleted. The ESSA User Manual is provided as Annex 4.

7 Overview of data types generated and collected in DRIVE

DRIVE will generate/collect data containing personal identifiers. Within personal data, we can

differentiate between individual level datasets and aggregated datasets. The individual level datasets should be sufficiently anonymized so there is no risk for re-identification using this data. These data may be primary data produced by consortium partners and associated partners or secondary data from existing registries or databases. All personal data will be generated within the activities of WP7.

The WP4 data management team will generate a meta-data repository of all datasets in collaboration with data owners and WP7 leads. This repository will be updated regularly.

8 Operational data management requirements for DRIVE research projects

Each individual study within DRIVE will need to provide a metadata on the used data, which will be added to this DMP.

The Innovative Medicines Initiative (IMI) requires a dataset template to be completed for each dataset described in the protocol or SAP of each study. This dataset template requires at the minimum the following:

- Dataset reference and name
- Dataset description
- Standards and metadata
- Data sharing policy
- Ethics and legal issues
- Data protection, IP/copyright and ownership

This dataset template shall be completed when the dataset is first imported and the whole set of completed dataset templates shall be reviewed every six months. The completed dataset templates will be enclosed in this DMP in Annex 5.

8.1 Requirements for metadata

All data custodians will be required to fill in the metadata template shown in Table 3 below for each dataset. These metadata are specifications for data that provide the contextual information required to understand those data. The template will be made available in the server. Each completed table will be reviewed by the data management team for completeness, compliance with the DMP and with the Consortium Agreement.

Table 3 Metadata requested per dataset (adapted from the Data Management General Guidance of the DMP Tool)[2]

General Overview	
Title	Name of the dataset or research project that produced it
DRIVE task	DRIVE task/subtask where dataset was generated
Data custodian	Names and addresses of the organizations

[2] https://dmptool.org/dm_guidance#metadata

	or people who own the data
Identifier	Unique number used to identify the data.
Start and end date	Study start and end date
Time period covered by the dataset	Start and end date of the period covered by the dataset
Methods	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data	Datasets containing personal data Datasets containing non-personal data
Processing	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description	
Subject	Keywords or phrases describing the subjects or content of the data
Language	All languages used in the dataset
Variable list and codebook	All variables in the data files, with description of the variable name, length, type, values
Data quality	Description of data quality standards and procedures to assure data quality
Technical description	
File inventory	All files associated with the project, including extensions
File formats	Format of the file
File structure	Organization of the data file(s) and layout of the variables, where applicable
Checksum	A digest value computed for each file that can be used to detect changes
Necessary software	Names of any special-purpose software packages required to create, view, analyse, or otherwise use the data
Access	
Rights	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Where and how your data can be accessed by other researchers
Data sharing	Description of how data will be shared, including access procedures

Ethics and legal issues	Description of any ethics and legal issues associated with the dataset, if any

8.2 Responsibilities of the data owner

Data owners per study and dataset will be identified in the dataset metadata. The data owner of the respective datasets must ensure and is responsible to comply with all legal and ethical requirements for data collection, handling, protection and storage. This includes adherence to regulations, guidelines such as (but not limited to) the EU clinical trial directive 2001/20/EC, Good clinical practice (GCP) and Good Pharmacoepidemiology Practice (GPP), as applicable.

9 Season 2018-2019 Dataset Overview

9.1 Dataset from study sites

Study Name	Data Source	Upload Date	Data Lock Date
Brand-specific influenza vaccine effectiveness in Europe, season 2018/19	National Institute for Health and Welfare - THL, FINLAND	14/05/2019	29/05/2019
	RCGP RSC, UNITED KINGDOM	17/06/2019	17/06/2019
	Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili, ITALY	28/05/2019	29/05/2019
	University of Athens, GREECE	27/05/2019	29/05/2019
	Helsinki University Hospital, FINLAND	17/05/2019	29/05/2019
	Vall d'Hebron, SPAIN	15/05/2019	29/05/2019
	National Institute for Infectious Diseases "Prof. Dr. Matei Balș", ROMANIA	21/05/2019	29/05/2019
	BIVE Hospital Network, ITALY	15/05/2019	29/05/2019
	Medical University Vienna, AUSTRIA	09/05/2019	29/05/2019

	Istituto Superiore di Sanita, ITALY	16/05/2019	29/05/2019
	FISABIO, SPAIN	16/05/2019	29/05/2019
	Influenza sentinel surveillance, Luxembourg	17/06/2019	17/06/2019

9.2 Resulting datasets generated on DRIVE Research Server

9.2.1 Season 2018-2019

As per the SAP, the following datasets are generated for each dataset specified in 8.1.1.

- Attrition diagram
- Table of outcomes by covariates (sex, age group, chronic condition, pregnancy, number of GP visits, number of hospitalizations, vaccination status previous year, vaccine brand) by influenza strain (A unspecified, A(H1N1), A(H3N2), B unspecified, B(Vict), B(Yam), Non-influenza, A, B, All)
- Table for data quality report (total number of subjects, number of subjects with influenza, number of vaccinated subjects)
- Histogram of covariates (see above) by controls/cases (for case-control studies) or exposed/unexposed persons (for cohort studies)
- Histogram of cumulative number of vaccinations over time
- Histogram of infections over time
- Pie chart with distribution of vaccine brands

Additional data quality reports have been generated for the data uploaded for each site, these reports can be found in Annex 2.

10 Sharing and secondary use of DRIVE generated or collected data

10.1 Procedures for making data findable

The information collected and updated via Table 3 will be available in the electronic study support web application. This will enable the easy identification of datasets available and identify the data owner.

10.2 Re-use of data within the DRIVE consortium

To achieve the objectives of DRIVE, it is imperative to follow the collaborative approach the partners agreed on when signing the consortium agreement. This includes the necessity to

share data from the individual studies for the implementation of the DRIVE project, while respecting data protection and intellectual property of the partners' work. For those individual studies within DRIVE that need to use data generated in another DRIVE task, the metadata will contain the data owner contact details to whom a requester can reach out if they need to access the results.

10.3 Re-use of DRIVE results by third parties

For those external individuals/institutions wanting to use DRIVE generated or collected data during the course of DRIVE, the Data Management Team should be contacted (Table 2). Given the nature of the studies conducted in DRIVE, that aim to make secondary use of data to conduct pooled data analysis, only access to pooled aggregated results datasets will be considered, since individual study data are not owned by DRIVE. Giving access to external parties will be considered by the Steering Committee on a case by case basis. Access rules for the time after DRIVE termination will be worked out and described in the final DMP.

11 Protection of personal data

The collection of personal data will be conducted under the applicable international, IMI, and national laws and regulations and requires previous written informed consent by the individual, i.e., with public and commercial entities and if applicable outside the EU in countries with lower data protection standards. To obtain the agreement of participants of studies to use their data for secondary research, the following lines can be included in the consent form:

- I understand the information collected about me will be stored in secure database, which will be used for future research.
- I authorise the research to use my anonymised study data for additional medical and/or scientific research projects.

The following points to consider will guide the protection of data within the DRIVE project:
(i) The entity providing personal data to the project shall verify that:

- the initial collection of these data has been compliant with the requirements of the original purpose
- the collection and the provision of the data to the project meets all legal requirements to which the entity is subject
- further storage and processing of the data after completion of the research project is in compliance with applicable law

(ii) The entity which provides personal data to the project shall document any restriction of use or obligation applicable to these data (e.g., the limited scope of purpose imposed by the consent form). The entity which uses personal data in the project shall be responsible to ensure that it has the right under the applicable data protection and other laws to perform the activities contemplated in the project.

DRIVE may also utilize purely register-based data that is collected as part of routine surveillance or clinical practice. Because of the nature of these data and the large number of study subjects, it is often not possible to obtain individual informed consent in these cases. However, all the other considerations related to ethics, data security and protection apply.

DRIVE researchers commit to the highest standards of data security and protection in order to preserve the personal rights and interests of study participants. They will adhere to the provisions set out in the:

- General data protection regulation (GDPR), into effect since May 2018[3]
- Directive 2006/24/EC of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communication services or of public communications networks[4]
- Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)[5]
- Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data[6]

To secure the confidentiality, accuracy, and security of data and data management, the following measures will be taken:

- All personal data obtained at the site level and transferred to the DRS for the DRIVE studies will be anonymized prior to the transfer. Keys to identification numbers will be held confidentially within the respective site. In situations were re-identification of study participants becomes necessary, for example the collection of additional data, this will only be possible through the sites and in cases where informed consent for such cases has been given.
- Personal data are entered through secure protocols. Data are processed only for the purposes outlined in the patient information and informed consent forms of the respective case studies. Use for other purposes will require explicit patient approval obtained through the respective site. Also, data are not transferred to any places outside the consortium without patient consent.
- None of the personal data will be used for commercial purposes, but the knowledge derived from the research using the personal data may be brought forward to such use as appropriate, and this process will be regulated by the Grant Agreement and the Consortium Agreement, in accordance with any generally valid legislation and regulations.

Personal data shall always be collected, stored, and exchanged in a secure manner, through secure channels.

Please also refer to the Research Collaborator Agreement (Annex 6), serving as atemplate data sharing agreement used for agreements between the DRIVE consortium and individual study sites.

[3] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0024&from=en
[4] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0024&from=en
[5] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002L0058&from=en
[6] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=en

12 Ethical aspects

DRIVE partners and associated partners are required to adhere to all relevant international, IMI, and national legislation and guidelines relating to the conduct of studies. Research ethics in DRIVE are described in detail in deliverable 6.2 (Report with the definition of the ethics policies handbook collection).

13 List of abbreviations

CA	Consortium Agreement
DMP	Data Management Plan
DoA	Description of Action
DRIVE	Development of Robust and Innovative Vaccine Effectiveness
EU	European Union
GDPR	General Data Protection Regulation
IMI	Innovative Medicines Initiative
SAP	Statistical Analysis Plan
WP	Work Package
ESSA	Electronic Study Support Application
sFTP	Secure File Transfer Protocol
SLA	Service Level Agreement
IP	Intellectual Property
MDR	Minimal Dataset Requirements
URS	Universal Routing Server
DRS	DRIVE Research Server

Annex 1. Data access request form

DRIVE Research Server application instructions

The DRIVE Research Server maintained by P95 CVBA in the DRIVE Research Consortium allows researchers to work together in a secure environment on studies developed in the DRIVE Research Consortium with the possibility to work with distributed data.

Access to the DRIVE Research Server is secured by two-factor authentication using Duo Access using a mobile phone running either the Android or iOS operating system or using a desktop application.

Please fill in the form and send it as a PDF by email to tom.desmedt@p-95.com (cc Roberto.bonauiti@unifi.it and turunen_top@gva.es) and add [DRIVE RESEARCH SERVER ACCESS REQUEST] to the subject heading, or send it by regular mail to the address below.

Tom De Smedt
Server Admin DRIVE Research Server
P95 CVBA
Koning Leopold III laan 1
3001 Heverlee
BELGIUM
Email: tom.desmedt@p-95.com
Tel: +32 (0) 472 21 65 65

DECLARATION REGARDING THE ACCESS TO THE DRIVE RESEARCH SERVER

The user requests access to the DRIVE Research Server using the two-factor authentication procedures set forward by the DRIVE Consortium in WP4.

Rationale to request access

User Responsibility Statement - The signatory hereby commits to use the access to the DRIVE Research Server and the accompanying two-factor authentication procedure only in the framework of the project and only to carry out its assigned tasks as described in the study protocol and related documents.

The signatory declares to adhere to the project study protocols and to be aware of the security measures for the use of the DRIVE Research Server. The signatory is aware that no data can be copied or used for other purposes. This includes a commitment not to transfer the access credentials to anyone else inside or outside his/her organisation and to take the necessary steps to prevent the access credentials from being misused in any way.

In case the signatory detects any use which is not compliant with the terms above he/she commits to inform **immediately** the DRIVE Server Admin by email to tom.desmedt@p-95.com

APPLICANT INFORMATION

First Name	***Last Name***
Organization	***Department***
Street	***Postal Code***
City	***Country***
Telephone Number	***Email Address***
Date of Application	***Signature***

Annex 2. SAP

Brand-specific influenza vaccine effectiveness in Europe Statistical Analysis Plan Season 2018/19

777363 - DRIVE

Development of robust and innovative vaccine effectiveness

WP7 - IVE studies

Lead contributor	Kaatje Bollaerts (P95)
Other contributors	Anke Stuurman (P95) Maria Alexandridou (P95) Marga Riera (P95)

Due date	NA
Delivery date	NA
Deliverable type	R
Dissemination level	PU

Document History

Version	Date	Contribution	Description
0.3	14.02.19	Kaat Bollaerts (P95), Anke Stuurman (P95), Maria Alexandridou (P95), Marga Riera (P95)	First, draft, shared with WP7
	21.02.19	Joan Puig Barbera (FISABIO), Topi Turunen (FISABIO), Caterina Rizzo (BIVE-HOSP), Stefania Bellino (ISS), Jorne Biccler (P95), Ulrike Baum (THL), Ritva Syrjänen (THL), Javier Diez-Domingo (FISABIO), Ainara Mira Iglesias (FISABIO)	WP7 review
	21.02.19	Monika Redlberger-Fritz (MUW), Caterina Rizzo (BIVE-HOSP), Stefania Bellino (ISS), Ulrike Baum (THL), Ritva Syrjänen (THL), Javier Diez-Domingo (FISABIO), Ainara Mira Iglesias (FISABIO), Magda Campins Marti (VHUH), Uy Hoang (UNIS), Daniela Pitigoi (NISS), Anca Drăgănescu (NISS), Oana Săndulescu (NISS), Helena Maltezou (UoA), Eva Martinez Ochoa (RS)	Input study sites
0.4	26.02.19	Kaat Bollaerts (P95), Maria Alexandridou (P95)	Addressing WP7 comments, shared with ISC for review
0.5	13.03.19	Kaat Bollaerts (P95), Maria Alexandridou (P95)	Addressing ISC comments, shared with EFPIA for review
0.6	17.04.19	Kaat Bollaerts (P95),	Addressing EFPIA comments, shared with ISC for endorsement

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of abbreviations

aTIV	Adjuvanted trivalent influenza vaccine
BIVE-HOSP	Italian Hospital Network
BMI	Body mass index
CI	Confidence interval
CIRI-IT	Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili
DRIVE	Development of Robust and Innovative Vaccine Effectiveness
DRIVE ESSA	DRIVE Electronic Study Support Application
DRIVE QCAC	DRIVE Quality Control and Audit Committee
ECDC	European Centre for Disease Prevention and Control
EMA	European Medicines Agency
FISABIO	Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana
GPP	Good Participatory Practice
HCW	Healthcare worker
HUCH	Helsinki University Central Hospital
ILI	Influenza like illness
IMI	Innovative Medicines Initiative
ISS	Istituto Superiore di Sanita
IVE	Influenza vaccine effectiveness
LAIV	Live-attenuated influenza vaccine
LCI	Laboratory-confirmed influenza
MUV	Medical University Vienna
NIID	National Institute for Infectius Diseases "Prof. Dr. Matei bals"
OR	Odds ratio
QIV	Non-adjuvanted quadrivalent influenza vaccine
RCGP	Royal College of General Practitioners
RE MA	Random-effects meta-analysis
REML	Restricted maximum likelihood
RR	Relative risk
RT-PCR	Reverse transcription polymerase chain reaction
SAP	Statistical Analysis Plan
SARI	Severe acute respiratory infection
THL	National Institute for Health and Welfare
TIV	Non-adjuvanted trivalent influenza vaccine
TND	Test negative design
UNIS	University of Surrey
UoA	National and Kapodistrian University of Athens
UK	United Kingdom
VE	Vaccine effectiveness
VHUH	Vall d'Hebron University Hospital

1 Background

The Development of Robust and Innovative Vaccine Effectiveness (DRIVE) project is a public-private partnership aiming to build capacity in Europe for estimating brand-specific influenza vaccine effectiveness (IVE). The DRIVE Project, which is funded by the Innovative Medicines Initiative (IMI), was initiated as a response to the changes for licensing of influenza vaccines in Europe. The new guidance on influenza vaccines by the European Medicines Agency (EMA) came into effect in the beginning of 2017. This guidance states that the performance of influenza vaccines should no longer be assessed based on serological assays, but should be based on post-authorization effectiveness studies [1].

In DRIVE, data from several independently operating national or regional study sites will be analysed jointly to obtain sufficient geographical coverage and sample size for brand-specific IVE estimates.

In 2017/18, a pilot study was performed to test the different operational aspects of the DRIVE project, including the IT infrastructure, the DRIVE governance for conducting IVE studies and to streamline key processes such as data collection, statistical analyses and dissemination of study results [2]. In the pilot study, there were four test-negative design studies (TND) and one register-based cohort study. The tools and processes developed during the pilot season 2017/18, will be used and further improved in the 2018/19 season.

The main objective of the 2018/19 season is to estimate brand-specific seasonal IVE in Europe by health care setting and age group. The DRIVE platform is still expanding, and not all vaccine brands used in Europe will be covered during the 2018/19 season.

This Statistical Analysis Plan (SAP) describes the characteristics of the participating study sites, the site-specific statistical analysis as well as the statistical analysis to pool data across study sites for the 2018/19 influenza season.

2 Reference documents

For this pilot season, the SAP has been developed using the following documents:

- DRIVE Generic protocols (D7.1 and D7.2)
- DRIVE 2018/19 local study protocols
- DRIVE Generic SAP: combining information on Influenza Vaccine Effectiveness across study sites (D4.4)
- DRIVE data management plan (D4.2)

The following supplementary files are provided:

- DRIVE minimal data requirements (ANNEX 1)
- DRIVE Electronic Study Support Application (ESSA) user manual (ANNEX 2)
- List of chronic conditions by study site

3 Study team

3.1 DRIVE WP7

Margarita Riera. MD. Epidemiologist. P95, Leuven, Belgium.
Kaatje Bollaerts PhD. Statistician. P95, Leuven, Belgium.
Anke Stuurman. MSc. Epidemiologist. P95, Leuven, Belgium.
Maria Alexandridou MSc. Data analyst. P95, Leuven, Belgium.
Jorne Biccler PhD. Statistician. P95, Leuven, Belgium.
Tom De Smedt Ir. DRIVE Research server system administrator. P95, Leuven, Belgium.
Topi Turunen. MD, Scientific Project Manager, FISABIO-Public Health.
Ritva Syrjänen. MD, PhD. Expert Scientist. Department of Public Health Solutions. THL, Tampere, Finland
Hanna Nohynek. MD PhD Chief Physician. Deputy Head of Unit of Infectious Diseases Control and Vaccinations. Co-lead of Expert group on Influenza and Viral Acute Respiratory Infections. THL, Helsinki Finland.
Miriam Levi. MD, PhD. Epidemiology Unit, Department of Prevention, Local Health Unit Tuscany Center, Florence, Italy.
Caterina Rizzo. MD. Epidemiologist. Ospedale Pediatrico Bambino Gesu (Rome, Italy). National Focal Point for the Influenza and other respiratory viruses ECDC Program.
Uy Hoang. Dr. MBBS, MPH, Research Fellow, University of Surrey, Guildford, UK.
Stefania Bellino. PhD. Statistician, Senior Researcher, Istituto Superiore di Sanità, Rome, Italy
Ornella Punzo. MD, MScPH, Internal Medicine Specialist, Researcher, Istituto Superiore di Sanità, Rome, Italy

3.2 Independent Scientific Committee (ISC)

Hector Izurieta, MD, Food & Drug Administration, United States

Liz Miller, Prof, Public Health England, United Kingdom

Mark Miller, MD, National Institutes of Health, United States

Marianne van der Sande, MD, Prof, Institute of tropical medicine Antwerp, Belgium

Stefania Salmaso, MD, independent consultant; formerly National Institute of Health, Italy

3.3 Study sites

Medical University Vienna (MUV), Austria

Monika Redlberger-Fritz. PhD, MD, Center of Virology, Medical University Vienna (MUW), Kinderspitalgasse 15, 1090 Vienna, Austria. Principal investigator, coordinating investigator and protocol author.

Therese Popow-Kraupp. Prof. MD, Center of Virology, MUW, Kinderspitalgasse 15, 1090 Vienna, Austria. Investigator and protocol author.

Helsinki University Central Hospital (HUCH), Finland

Kirsi Skogberg. MD PhD Specialist in infectious diseases and internal medicine. Division of Infectious Diseases, HUCH Inflammation Center, Helsinki University Central Hospital, Jorvi Hospital, Espoo, Finland. Primary investigator.

Raija Auvinen. MD, clinical coordinator, Division of Internal Medicine and Rehabilitation, Helsinki University Hospital, Finland. Co-investigator.

Outi Debnam. Division of Infectious Diseases, HUCH Inflammation Center, Helsinki University Hospital, Finland. Study nurse.

Marja-Leena Michelsson. Division of Infectious Diseases, HUCH Inflammation Center, Helsinki University Hospital, Finland. Study nurse (part time).

Raisa Loginov. PhD, specialist in virology, Helsinki University Hospital, Helsinki. Co-investigator.

Niina Ikonen. MSc, Senior specialist (virology), Department of Health Security, Unit of Expert Microbiology, Helsinki. Co-investigator.

Ritva Syrjänen. MD, PhD, Senior researcher, Department of Public Health Solutions, Public Health Evaluation and Projection Unit, Tampere. Co-investigator.

The National Institute for Health and Welfare (THL), Finland

Hanna Nohynek. MD PhD Chief Physician. Deputy Head of Unit of Infectious Diseases Control and Vaccinations. Co-lead of Expert group on Influenza and Viral Acute Respiratory Infections. National Institute for Health and Welfare (THL), Helsinki Finland. Principal Investigator of the study.

Ulrike Baum. MSc. Statistical researcher. Department of Public Health Solutions. THL, Helsinki, Finland. Principal statistician of the study.

Ritva Syrjänen. MD PhD Senior Researcher. Department of Public Health Solutions. THL, Tampere, Finland. Clinical trials and register study specialist of the study.

Pregnancy: 1st Department of Obstetrics and Gynecology, "Alexandra" General Hospital of Athens, National and Kapodistrian University of Athens (UoA), Medical School, Athens, Greece

Helena Maltezou. Dr. Head, Department for Interventions in Healthcare Facilities, Hellenic Center for Disease Control and Prevention, Athens, Greece, Principal investigator.

Alexandros Rodolakis. Prof. 1st Department of Obstetrics and Gynecology, National and Kapodistrian University of Athens, «Alexandra» General Hospital of Athens, Athens, Greece. Co-principal investigator.

Dimitrios Loutradis. Prof. , Chairman 1st Department of Obstetrics and Gynecology, National and Kapodistrian University of Athens, «Alexandra» General Hospital of Athens, Athens, Greece. Scientific advisor.

Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI-IT), Italy

Giancarlo Icardi. MD Director of CIRI-IT. Full Professor of Hygiene and Preventive Medicine Department of Health Sciences, University of Genoa, Italy. Hospital Policlinico San Martino Genoa. Principal investigator. Laboratory Coordinator.

Donatella Panatto. PhD. Member of the CIRI-IT scientific committee. Associate Professor of Hygiene and Preventive Medicine Department of Health Sciences, University of Genoa, Italy. Co-investigator.

Paolo Durando. MD. PhD. Member of the CIRI-IT scientific committee. Associate Professor of Occupational Medicine Department of Health Sciences, University of Genoa, Italy. Hospital Policlinico San Martino Genoa. Co-investigator.

Andrea Orsi. MD. PhD. Member of the CIRI-IT scientific committee. Researcher of Hygiene and Preventive Medicine Department of Health Sciences, University of Genoa, Italy. Hospital Policlinico San Martino Genoa. Co-investigator. Laboratory Coordinator of Genoa Unit

Piero Luigi Lai. PhD. Department of Health Sciences, University of Genoa, Italy. Technical project and data manager.

Elena Pariani. PhD. Member of the CIRI-IT scientific committee. Department of Biomedical Sciences for Health, University of Milan, Italy. Co-investigator. Laboratory Coordinator of Milan Unit

Silvana Castaldi. Department of Biomedical Science for Health, University of Milan, Italy - Fondazione IRCCS Ca' Granda Ospedale Maggiore di Milano, Italy. Co-investigator.

Istituto Superiore di Sanita (ISS), Italy

Maria Rita Castrucci. PhD. Biologist. Head of the National Influenza Center at the Department of Infectious Diseases, ISS, Rome, Italy. Influenza virus expert.

Simona Puzelli. PhD. Biologist at the Department of Infectious Diseases, ISS, Rome, Italy. Influenza virus expert.

Antonino Bella. Statistician in the Department of Infectious Diseases of the National Institute of Health, ISS, Rome, Italy. Head of the Influenza Surveillance System (InfluNet). Statistician and data manager.

Italian Hospital Network (IT-BIVE-HOSP), Italy

Caterina Rizzo. MD. Innovation Unit and Clinical Paths, Directorate of Clinical Departments. Bambino Gesù Children's Hospital, Rome, Italy. Principal investigator.

Elisabetta Pandolfi. MD Innovation Unit and Clinical Paths, Directorate of Clinical Departments. Bambino Gesù Children's Hospital, Rome, Italy. Technical project manager.

Francesco Gesualdo. MD Innovation Unit and Clinical Paths, Directorate of Clinical Departments. Bambino Gesù Children's Hospital, Rome, Italy. Technical project manager.

Christian Napoli. MD Department of medical-surgical sciences and translational medicine, University of Rome "Sapienza", Lazio Region. Principal Investigator in one of the participating hospitals.

Maria Chironna. Biologist. Department of Biomedical science and medical Oncology of the University of Bari, Puglia Region. Principal Investigator in one of the participating hospitals.

Andrea Orsi. MD IRCCS University Hospital San Martino, Genoa, Liguria Region. Principal Investigator in one of the participating hospitals.

Ilaria Manini. Biologist. Department of Physiopathology, Experimental Medicine and Public Health, University of Siena, Tuscany Region Principal Investigator in one of the participating hospitals.

National Institute for Infectious Diseases "Prof. Dr. Matei Balș", Romania

Anca Drăgănescu, MD, PhD. Senior specialist paediatrics. Specialist infectious diseases. National Institute for Infectious Diseases "Prof. Dr. Matei Balș", Bucharest, Romania. Project manager and principal investigator – children.

Daniela Pițigoi, MD, PhD. Senior specialist epidemiology. Assoc. Prof. Epidemiology. National Institute for Infectious Diseases "Prof. Dr. Matei Balș", Bucharest, Romania. Technical project manager.

Oana Săndulescu, MD, PhD. Specialist infectious diseases. Assoc. Prof. infectious diseases National Institute for Infectious Diseases "Prof. Dr. Matei Balș", Bucharest, Romania. Principal investigator – adults.

Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO), Spain

Javier Díez-Domingo. PhD. MD. Paediatrician. Head of the Vaccine Research Department of FISABIO-Public Health. (Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana). Avenida de Cataluña, 21, 46020, Valencia. Principal investigator.

Ainara Mira-Iglesias. Mathematician. Master's degree in Biostatistics. Statistician in the Vaccine Research Department of FISABIO-Public Health. Avenida de Cataluña, 21, 46020, Valencia. Data analyst of the study.

F. Xavier López-Labrador. PhD. Virologist. Senior Laboratory Coordinator in the Genomics Department of FISABIO-Public Health. Avenida de Cataluña, 21, 46020, Valencia. Consorcio de Investigación Biomédica de Epidemiología y Salud Pública (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain. Laboratory Coordinator and Virologist of the study.

Javier García-Rubio. Pharmacist. Master's degree in Clinical Trials. Research assistant in the Vaccine Research Department of FISABIO-Public Health. Avenida de Cataluña, 21, 46020, Valencia. Data Manager of the study.

Mario Carballido-Fernández. PhD. MD. Hospital General Universitario de Castellón, Castellón, Spain. Universidad CEU Cardenal Herrera, Castellón, Spain. Principal investigator in one of the participating hospitals.

Miguel Tortajada-Girbés. PhD. MD. Hospital Universitario Doctor Peset, Valencia, Spain. Principal investigator in one of the participating hospitals.

Juan Mollar-Maseres. PhD. MD. Hospital Universitario y Politécnico La Fe, Valencia, Spain. Principal investigator in one of the participating hospitals.

German Schwarz-Chavarri. PhD. MD. Hospital General Universitario de Alicante, Alicante, Spain. Principal investigator in one of the participating hospitals**.**

Vall d'Hebron University Hospital (VHUH), Barcelona, Spain

Magda Campins Martí. (MD,PhD) Head of the Department, Preventive Medicine and Epidemiology Department, HUVH, Barcelona, Spain. Principal investigator.

José Ángel Rodrigo Pendás (MD) Clinical researcher, Preventive Medicine and Epidemiology Department, HUVH, Barcelona, Spain. Investigator.

Sonia Uriona (MD) Clinical researcher, Preventive Medicine and Epidemiology Department, HUVH, Barcelona, Spain. Investigator.

Eva del Amo (PhD) Researcher, Preventive Medicine and Epidemiology Department, HUVH, Barcelona, Spain. Research technician.

Andrés Antón Pagarolas (PhD) Head of Respiratory Viruses Unit, Microbiology Department, HUVH, Barcelona, Spain. Investigator.

Royal College of General Practitioners (RCGP) & University of Surrey (UNIS), United Kingdom

Prof Simon de Lusignan. Prof. Professor of Primary Care and Clinical Informatics. Chair in Health Care Management, University of Surrey, UK. Principal Investigator.

Uy Hoang. Research Fellow. Department of Clinical and Experimental Medicine, University of Surrey, UK.

Ivelina Yonova. Department of Clinical and Experimental Medicine, University of Surrey, UK. Practice Liaison Officer.

Rachel Byford Senior Database Developer. Department of Clinical and Experimental Medicine, University of Surrey, UK.

Dr Mark Joy. Senior Lecturer in Data Modelling and Population Health. Department of Clinical and Experimental Medicine, University of Surrey.

Filipa Ferreira. PhD Department of Clinical and Experimental Medicine, University of Surrey, UK. Project Manager

Tristan Clark. Associate Professor and Honorary Consultant in Infectious Diseases. NIHR Post-Doctoral Fellow. Academic Unit of Clinical and Experimental Sciences, University of Southampton.

4 Objectives

4.1 Primary objective

To estimate seasonal **overall** and **brand-specific** IVE against laboratory-confirmed influenza stratified by setting (primary care, hospital-based or mixed setting in case the source of the cases cannot be obtained) and age group (6m-17yr, 18-64 yr, ≥ 65yr), by type of outcome:

- any laboratory-confirmed influenza
- laboratory-confirmed influenza A, overall and by subtype (A(H1N1), A(H3N2))
- laboratory-confirmed influenza B, overall and by lineage (B Victoria, B Yamagata)
- brand-specific IVE only: any laboratory-confirmed influenza subtype/lineage included in the vaccine brand

4.2 Secondary objective

To estimate seasonal **vaccine-type** IVE against laboratory-confirmed influenza stratified by setting (primary care, hospital-based or mixed) and age group (6m-17yr, 18-64 yr, ≥65yr), by type of outcome:

- any laboratory-confirmed influenza
- laboratory-confirmed influenza A, overall and by subtype (A(H1N1), A(H3N2))
- laboratory-confirmed influenza B, overall and by lineage (B Victoria, B Yamagata)

any laboratory-confirmed influenza subtype/lineage included in the vaccine type
The following vaccine types will be considered:

- Trivalent non-adjuvanted
- Trivalent adjuvanted
- Quadrivalent live attenuated
- Quadrivalent inactivated

4.3 Exploratory objective

To estimate seasonal **overall** IVE against laboratory-confirmed influenza stratified by setting (primary care, hospital-based or mixed) and age group (6m-17yr, 18-64 yr, ≥65yr whenever relevant),within **risk groups**, by type of outcome:

- any laboratory-confirmed influenza
- laboratory-confirmed influenza A, overall and by subtype (A(H1N1), A(H3N2))
- laboratory-confirmed influenza B, overall and by lineage (B Victoria, B Yamagata)

The following risk groups will be considered:

- Pregnant women
- Healthcare workers
- Presence of chronic conditions by the following sub-categories;
    - Cardiovascular disease
    - Lung disease
    - Diabetes

Pregnant women and healthcare workers were selected as risk groups of interest as two studies were specifically designed to investigate these risk groups (pregnancy study by University of Athens, healthcare workers study by CIRI-IT, Italy). The three chronic conditions (cardiovascular disease, lung disease and diabetes) were chosen to explore the feasibility of estimating IVE by risk group as they are believed to be chronic conditions with the highest prevalence (see Section 12.4).

5 Study design

A multi-centre study with data available from four primary care based TND studies, six hospital based TND studies, one register-based cohort and two clinical cohorts (in pregnant women and their young infants and in healthcare workers). A list of the participating study sites according to study design and setting is given in Table 1. All the TND studies and the register-based cohort follow closely the DRIVE generic protocols (D7.1 and D7.2) for their respective study designs. The characteristics of the site-specific studies are summarized in Table 2 for the TND studies and Table 3-Table 5 for the cohort studies. More details on the individual studies are provided in the subsequent sections. When feasible, additional site-specific studies might be included in the analysis if their data becomes available prior to 15th May 2019.

National Institute for Infectious Diseases "Prof. Dr. Matei Balș", Romania

Table 1. Overview of the participating study-sites, 2018/19

Test-negative design studies, primary care: 1. Medical University Vienna (MUV), Austria 2. Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI-IT), Italy 3. Royal College of General Practitioners (RCGP) & University of Surrey (UNIS), United Kingdom 4. Istituto Superiore di Sanita (ISS), Italy
Test-negative design studies, hospital based: 1. Medical University Vienna (MUV), Austria 2. Helsinki University Central Hospital (HUCH), Finland 3. Italian Hospital Network (IT-BIVE-HOSP), Italy 4. Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI-IT), Italy 5. Vall d'Hebron University Hospital (VHUH), Barcelona, Spain 6. Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO), Spain 7. National Institute for Infectious Diseases 'Prof. Dr. Matei Bals' (NIID), Romania
Register-based cohort study: 1. The National Institute for Health and Welfare (THL), Finland
Clinical cohort studies: 1. Pregnancy: 1st Department of Obstetrics and Gynecology, "Alexandra" General Hospital of Athens, National and Kapodistrian University of Athens (UoA), Medical School, Athens, Greece 2. Healthcare workers: Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI-IT), Italy

Table 2. Overview of test-negative design study sites characteristics - 2018/19

Site	MUV	CIRI-IT	ISS	RCGP/ UNIS	MUV	HUCH	BIVE-HOSP	NIID	VHUH	FISABIO
Country	Austria	Italy	Italy	UK	Austria	Finland	Italy	Romania	Spain	Spain
Setting	Primary care	Primary care	Primary care	Primary care	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital
Source of cases	80 primary care physicians	21 primary care physicians	Ca. 1000 primary care physicians	6 primary care practices	1 hospital	1 hospital	5 hospitals	1 hospital	1 hospital	4 hospitals
Population	General population ≥6 months	General population ≥6 months	General population ≥6 months	General population ≥6 months	General population ≥6 months	General population ≥18 years	General population ≥6 months	General population ≥6 months	General population ≥6 months	General population ≥6 months
Expected sample size (number lab confirmed)	900* (n.a.) *together with hospital	1,500 (500)	2,380 (n.a.)	1,200 (400)	900* (n.a.) *together with hospital	600 (125)	2,488 (n.a.)	400 (150)	1,600 (800)	2,000 (n.a.)
Start data collection	01.10.2018	05.11.2018	15.10.2018	11.02.2019	01.10.2018	26.11.2018	26.11.2018	12.11.2018	13.12.2018	10.09.2018
Case definition	ILI[(1)]	ILI[(1)]	ILI[(1)]	ILI[(1)]	SARI[(2)]	SARI[(2)]	SARI[(2)]	SARI[(2)]	SARI[(2)]	<5y:Hospitalized for any acute reason ≥5y: ILI[(3)]
Case										
Sampling strategy[(4)]	Undefined	Predefined rules	All	Predefined rules	Undefined	All	All	All	All	All
Type of swab	Naso-pharyngeal	Nasal or oropharyngeal	Throat swab	Nasal	Nasopharyngeal	Nasal and throat or nasopharyngeal	Pharyngeal or nasopharyngeal	<14y: nasopharyngeal and nasal ≥14y: nasopharyngeal and pharyngeal	< 18y: usually nasopharyngeal >18 y: nasopharyngeal and/or pharyngeal and/or bronchoalveolar l	<14y: nasopharyngeal and nasal ≥14y: nasopharyngeal and pharyngeal
Who swabs	HCW	HCW	HCW	HCW	HCW	HCW	HCW	HCW	HCW	HCW

Table 2. Overview of test-negative design study sites characteristics - 2018/19, continued

Site	MUV	CIRI-IT	ISS	RCGP/ UNIS	MUV	HUCH	BIVE-HOSP	NIID	VHUH	FISABIO
Country	Austria	Italy	Italy	UK	Austria	Finland	Italy	Romania	Spain	Spain
Laboratory test influenza	RT-PCR	RT-PCR or rapid diagnostic test	RT-PCR	RT-PCR desktop analyser	RT-PCR	RT-PCR	RT-PCR	RT-PCR	< 18y: Antigen detection > 18y: PCR	RT-PCR
A/subtype	Yes	Yes	Yes	No	Yes	Yes	Yes	Partial (H only)	Yes	Yes
B/lineage	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Laboratory test subtyping	RT-PCR	RT-PCR	RT-PCR	n/a	RT-PCR	Real-time RT-PCR	RT-PCR	RT-PCR		RT-PCR
Source of vaccination status	-Medical records -Otherwise, patient/ relatives interview	-Vaccine register -Medical records	-Medical records	- Medical records	-Medical records	-Vaccine register -Vaccine card	-Primary care physician interview (for patients indicating being vaccinated or not knowing vaccination status)	-Vaccine card -HCW interview -Patient/ relatives interview	-Vaccine register -Medical records -Vaccine card -Otherwise, patient/ relatives interview	-Vaccine register

Table 2. Overview of test-negative design study sites characteristics - 2018/19, continued

Site	MUV	CIRI-IT	ISS	RCGP/ UNIS	MUV	HUCH	BIVE-HOSP	NIID	VHUH	FISABIO
Country	Austria	Italy	Italy	UK	Austria	Finland	Italy	Romania	Spain	Spain
Covariates available for adjustment	Age, sex, date of swab, 1+ chronic condition, pregnancy	Age, sex, date of swab, 1+ chronic condition, pregnancy, nr of primary care visits in last 12 months, influenza vaccination in previous season	Age, sex, date of swab, 1+ chronic condition, nr of primary care visits in last 12 months, influenza vaccination in previous season	Age, sex, date of swab, 1+ chronic condition, pregnancy, nr of primary care visits in last 12 months, influenza vaccination in previous season	Age, sex, date of swab, 1+ chronic condition, pregnancy,	Age, sex, date of swab, 1+ chronic condition, pregnancy, nr of hospitalisations in last 12 months, influenza vaccination in previous season	Age, sex, date of swab, 1+ chronic condition, influenza vaccination in previous season, nr of hospitalisations in last 12 months, for 65+: frailty	Age, sex, date of swab, 1 chronic condition or more pregnancy, nr of hospitalisations in last 12 months, influenza vaccination in previous season	Age, sex, date of swab, 1+ chronic condition, pregnancy, nr of hospitalisations in last 12 months, influenza vaccination in previous season	Age, sex, date of swab 1+ chronic condition, pregnancy, nr of hospitalisations in last 12 months, influenza vaccination in previous season

H: hemagglutinin; ILI: influenza-like illness; LCI: laboratory-confirmed influenza; n/a: not applicable; HCW: healthcare worker; RT-PCR: Reverse transcription polymerase chain reaction. (1) ECDC case definition (2) IMOVE+ 2017/2018 case definition
(3) ECDC case definition, without "sudden onset" (4) Sampling strategies 1) all: all patients with ILI or SARI are sampled; 2) 'predefined rules': systematic sampling according to predefined rules; 3) 'undefined': non-systematic sampling at practioner's discretion. More details on the sampling strategies are given in Table 8.

Table 3. Overview of register-based cohort study, 2018/19

Site	**THL**
Country	Finland
Setting	Primary care and hospital
Source of cases	All healthcare facilities in Finland
Population	General population 6-months-6 years and ≥65 years
Population size	~1555300 (31.12.2017)
Start data collection	Ongoing
Case	LCI positive
Sampling strategy(1)	undefined
Type of swab	Nasopharyngeal swabs or nasal and/or throat swabs or nasopharyngeal aspirates (sometimes other clinical samples) analysed by real time RT-PCR, multiplex RT-PCR, culture and/or antigen detection
Who takes swab	HCW
Laboratory test influenza diagnosis	RT-PCR, Antigen detection
A/subtype available	No
B/lineage available	No
Laboratory test subtyping	n/a
Source of vaccination status	Vaccine register
Covariates available for adjustment	Age, sex, calendar week at influenza test, 1 chronic condition or more, number of hospitalizations in the last 12 months, number of primary care consultations in the last 12 months, influenza vaccination in previous season

LCI: laboratory-confirmed influenza; n/a: not applicable; RT-PCR: Reverse transcription polymerase chain reaction, HCW: healthcare worker
(1) Sampling strategies 1) all: all patients with ILI or SARI are sampled; 2) 'predefined rules': systematic sampling according to predefined rules; 3) 'undefined': non-systematic sampling

Table 4. Overview of clinical cohort study in pregnant women and their young adults, 2018/19

Site	**UoA**
Country	Greece
Source of cases	Department of obstetrics and gynaecology from 1 hospital
Population	Pregnant women (18y-≤45y), and their infants (≤6 months)
Targeted study size	700 pregnant women in the cohort, 25 laboratory confirmed influenza cases in pregnant women and 140 in infants
Start data collection	17.10.2018
Case definition	ILI (ECDC case definition)
Case	Above clinical case definition + LCI positive
Sampling strategy(1)	All
Type of Swab	Nasal-pharyngeal
Who takes swab	HCW
Laboratory test influenza diagnosis	PCR
A/subtype available	Yes
B/lineage available	Yes
Laboratory test subtyping	PCR
Source of vaccination status	Medical records
Covariates available for adjustment	Age, sex, date of swab, 1 chronic condition or more, gestational age at vaccination (for pregnant women), influenza vaccination in previous season, education, ethnicity, nr of household members, number of children < 5 years, number of labors in the past

ILI: influenza-like illness; LCI: laboratory-confirmed influenza; HCW: health care worker, RT-PCR: Reverse transcription polymerase chain reaction
(1) Sampling strategies 1) all: all patients with ILI or SARI are sampled; 2) 'predefined rules': systematic sampling according to predefined rules; 3) 'undefined': non-systematic sampling at practioner's discretion

Table 5. Overview of clinical cohort study in healthcare workers, 2018/19

Site	**CIRI-IT**
Country	Italy
Source of cases	2 hospitals
Population	Healthcare workers, all ages (≥ 18 years)
Targeted study size	6,000 health care workers in the cohort, 500-700 ILI cases and 200-400 laboratory confirmed influenza cases
Start data collection	08.10.2018
Case definition	ILI (ECDC case definition)
Case	Above clinical case definition + LCI positive
Sampling strategy	All
Swab	Nasal or oropharyngeal
Who takes swab	Self-collected or collected by CIRI-IT medical staff
Laboratory test influenza diagnosis	RT-PCR or rapid diagnostic test
A/subtype available	Yes
B/lineage available	Yes
Laboratory test subtyping	RT-PCR
Source of vaccination status	Vaccine register
Covariates for adjustment	age, sex, date at swab, 1 chronic condition or more, pregnancy, number of hospitalizations in the last 12 months, influenza vaccination in previous season

ECDC: European Center for Disease Prevention and Control; H: hemagglutinin; ILI: influenza-like illness; LCI: laboratory-confirmed influenza; RT-PCR: Reverse transcription polymerase chain reaction; y: years
*Sampling strategies 1) all: all patients with ILI or SARI are sampled; 2) 'predefined rules': systematic sampling according to predefined rules; 3) 'undefined': non-systematic sampling at practioner's discretion
systematic sampling according to predefined rules; 3) 'undefined': non-systematic sampling at practioner's discretion

6 Study population

In all TND studies and the register-based study, the population under study is the general population. In the two clinical cohort studies, the populations under study were pregnant women and their young infants and healthcare workers.

7 Study period

The start of the data collection for the 2018/19 influenza season differs between the sites (Table 2-Table 5).

For the TND studies, the study period for the analysis will start when the influenza virus circulation begins (first week of two consecutive weeks when influenza viruses are detected at the study site level, based on the data as provided to DRIVE) in the country/region and will finish after the influenza season (defined as the end of the week prior to the first of two consecutive weeks when no influenza viruses are detected at the study site level, based on the data as provided to DRIVE). The study period of analysis might be different for different study sites.

In the particular case of THL (Finland), data is continuously collected throughout the year since they use the national registers. The study period for analysis goes from week 40 till week 20.

8 Case definitions

8.1 Influenza-like illness (ILI)

A case of influenza like illness (ILI) will be defined by the ECDC case definition [3] as an individual who presents with a

- sudden onset of symptoms

AND, at least one of the following four systemic symptoms:

- fever or feverishness;
- malaise;
- headache;
- myalgia;

AND, at least one of the following three respiratory symptoms:

- cough;
- sore throat;
- shortness of breath.

8.2 Severe acute respiratory infection (SARI)

A case of severe acute respiratory infection (SARI) will be defined by the IMOVE+ 2017/2018 case definition as a hospitalised person, with at least one of the following systemic symptoms or signs;

- fever or feverishness;
- malaise;
- headache;
- myalgia;
- deterioration of general condition (asthenia or loss of weight or anorexia or confusion or dizziness)

AND at least one respiratory symptom or sign e.g.

- cough;
- sore throat;
- shortness of breath;

at admission or within 48 hours after admission.

The symptoms should not have started (or, if chronic, clearly worsened) more than 7 days before swabbing.

8.3 Adherence to the case definitions

All study sites follow the ILI or SARI clinical case definitions with the exception of FISABIO (Spain)

FISABIO (Spain, TND hospital-based): For children <5 years, a clinical case is defined as a person with a hospitalization for any acute reason whose symptom onset (of any symptom possibly related to influenza – Table 6) was in the 7 days prior to admission. For subjects 5 years and above, a modified ECDC ILI case definition is used, being hospitalized with at least one systematic symptom (fever or feverishness, malaise, headache or myalgia) and at least one respiratory symptom (cough, sore throat or shortness of breath) whose onset was in the 7 days prior to admission.

Table 6. FISABIO: symptoms possibly related to influenza

Eligibility diagnosis, symptoms and signs
Acute upper and lower respiratory disease
Dyspnea breathing anomaly, shortness of breath, tachypnea
Asthma
Pneumonia and influenza
Heart failure
Myalgia
Altered consciousness, convulsions, febrile convulsions
Fever or fever unknown origin or non specified
Cough
Apnea
Gastrointestinal manifestations
Sepsis, systemic inflammatory response syndrome

9 In- and exclusion criteria

9.1 Test-negative design studies

9.1.1 Recommended exclusion criteria

The following exclusion criteria will be applied to subjects presenting with ILI;

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. is institutionalised at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after ILI onset

6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation

The following exclusion criteria will be applied to subjects presenting with SARI;

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. is institutionalised at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after SARI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation
7. was previously hospitalised < 48 hours prior to SARI onset
8. had his/her ILI/SARI onset ≥ 48 hours after hospital admission

Note: a patient can be enrolled several times as long as he/she does not have a previous laboratory confirmed influenza for the current season.

9.1.2 Adherence to the recommended ILI/SARI exclusion criteria

An overview of the adherence to the ILI and SARI exclusion criteria at study recruitment is given in Table 7. All variables related to the exclusion criteria are listed as obligatory variables in the Minimal Data Requirements (ANNEX 1). Records that violate the exclusion criteria will be discarded at analysis stage, whenever possible.

Table 7. Test-negative design studies: overview of exclusion criteria applied at study recruitment, 2018/19

Site	MUV	MUV	HUCH	BIVE-HOSP	CIRI-IT	ISS	NIID	FISABIO	VHUH	RCGP/ UNIS
Country	Austria	Austria	Finland	Italy	Italy	Italy	Romania	Spain	Spain	UK
Setting	PC	HOSP	HOSP	HOSP	PC	PC	HOSP	HOSP	HOSP	PC
Clinical case definition	ILI	SARI	SARI	SARI	ILI	ILI	SARI	ILI	SARI	ILI
1. Unwilling or unable to give consent	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No(1)	Yes
2. Age <6 months at symptom onset	Yes	Yes	n/a	Yes	Yes	Yes	Yes	No*	Yes	Yes
3. Contraindication	No	Yes	Yes	Yes	Yes	Yes	Yes	No*	Yes	Yes*
4. Institutionalized	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes*
5. Respiratory specimen taken ≥ 8 days after ILI onset	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6. Prior influenza infection in current season	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes*
7. Previously hospitalised < 48 hours prior to ILI onset	n/a	Yes	Yes	Yes	n/a	n/a	Yes	Yes(2)	Yes	n/a
8. ILI onset ≥ 48 hours after hospital admission	n/a	Yes	Yes	Yes	n/a	n/a	Yes	Yes	Yes	n/a
Other local exclusion criteria	No	No	Yes(3)	Yes(4)	No	No	No	Ye(5)	Yes(6)	No

n/a: not applicable, ILI: influenza like illness

* Can be excluded at analysis stage

(1) No informed consent was required as no intervention required for the study fall outside the usual practice of the Hospital Universitari Vall d'Hebron during the influenza season. (2) Patients hospitalized < 30 days from the current hospitalisation are excluded. (3) Not a resident of Espoo, Kauniainen or Kirkkonummi. (4) Antiviral therapy; Remain

in hospital for less than 24 hours.(5) Not residing in hospitals catchment area for at least previous 6 months; Remains in hospital for less than 24 hours. (6) A patient not belonging to the Institut Català de la Salut network

9.2 Cohort studies

9.2.1 THL Finland: register-based cohort study

In the Finnish register-based cohort study, all subjects belonging to the study population and contributing data to the study period (starting 2018, week 40) are included, with the following exclusion criterion applied;

Exclusion criteria:

- subjects with presumably incomplete vaccination records in 2018/19 or 2017/18

9.2.2 Pregnancy cohort

The following in- and exclusion criteria will be applied to all study subjects;

Inclusion criteria:

- age 18 to ≤45 years
- stable health
- presented to the outpatient clinic of the department of obstetrics and gynaecology between October 1 and December 31, 2018

Exclusion criteria:

- is unwilling to participate or unable to communicate (in Greek or English) and give consent
- received influenza vaccine < 6 months prior to study entry
- received any investigational drug or product < 30 days prior to study entry
- history of Guillain-Barré syndrome
- history of hypersensitivity to influenza vaccines or its components
- immunosuppression
- received immunoglobulins or blood products < 3 months

9.2.3 Healthcare workers

The following in- and exclusion criteria will be applied to all study subjects;

Inclusion criteria:

- in service prior to start of follow-up in Week 42 2018

Exclusion criteria:

- is unwilling to participate or unable to communicate and give consent

10 Outcome

10.1 Outcome definition

The outcome of interest is laboratory-confirmed influenza, using the following definitions:

Estimating seasonal overall, brand-specific and type-specific IVE against **any** medically attended laboratory-confirmed **influenza** (stratified by healthcare setting and age group);

- Positive: any laboratory-confirmed influenza
- Negative: no laboratory-confirmed influenza

Estimating seasonal overall, brand-specific and type-specific IVE against any medically attended laboratory-confirmed **influenza type, subtype or lineage** (stratified by healthcare setting and age group);

- Positive: laboratory-confirmed influenza of the specific type, subtype or lineage of interest
- Negative: no laboratory-confirmed influenza

For trivalent vaccines, estimating seasonal brand-specific and type-specific IVE against any medically attended laboratory-confirmed influenza included in the vaccine

- Positive: laboratory-confirmed influenza of any of the subtypes and lineage included in the vaccine
- Negative: no laboratory-confirmed influenza

10.2 Case identification

For the TND studies, ILI and SARI cases are identified among all patients presenting to primary care or hospital.

At UoA (Greece, pregnancy cohort), all enrolled women are actively followed-up through weekly telephone calls asking about the onset of a febrile episode, acute respiratory infection, ILI, acute otitis media and/or pneumonia, SARI, healthcare seeking, hospitalization and use of antibiotics in women and their infants.

At CIRI-IT (Italy, HCW cohort), all participants were regularly sent reminders through e-mail to call the study team in case of ILI.

At THL (Finland, register-based cohort study), only positive results of the influenza tests are available.

10.3 Swab sampling strategy

Different sampling strategies were used for collecting respiratory samples from patients meeting the ILI/SARI clinical case definitions;

- 'all': all patients with ILI or SARI are sampled
- 'predefined rules': systematic sampling according to predefined rules
- 'undefined': non-systematic sampling at practitioner's discretion

The sampling strategies are different for the different sites and might also differ between subpopulations from the same study site. Details on the sampling strategies are given in Table 8.
Swabs are performed by HCW in all studies with the exception of the CIRI-IT HCW cohort, where swabs are self-collected or collected by CIRI-IT medical staff. Self-collected swabs have similar sensitivity to those taken by health-care workers [4, 5] and the extent of postal delay is not associated with the likelihood of PCR positivity for influenza [5].
The type of swabs are either nasal, nasopharyngeal, oropharyngeal, pharyngeal or throat swabs (Table 2-Table 5).
Samples taken >=8 days after ILI onset will be excluded from analysis.

Table 8. Test-negative design studies: overview of swab sampling strategies used, 2018/19

Sampling strategy	Site (Country)
All	BIVE-HOSP (Italy), HUCH (Finland), NIID (Romania) VHUH (Spain), FISABIO (Spain), UoA (Greece, pregnancy cohort), CIRI-IT (Italy, HCW cohort)
Predefined rules	CIRI-IT (Italy): Systematic sampling is encouraged, for example, the first 3 ILI that present each week ISS (Italy): Systematic sampling of the first 2 ILI patients that present each week, and if possible all ≥65years ILI cases RCGP/UNIS (UK): All cases of ILI are encouraged to be swabbed in this study, up to a maximum of 10 per practice, per day
Undefined rules	MUV (Austria), THL (Finland)

10.4 Laboratory testing

The influenza laboratory confirmation was done using antigen detection, culture, PCR, rapid diagnostic tests, or real-time RT-PCR, and subtyping/lineage testing was done using PCR or real-time-PCR. Except THL (Finland, register-based cohort) and RCGP/UNIS (UK), all sites are collecting information on influenza subtypes/lineages (A/H1N1, A/H3N2, B/Victoria, and B/Yamagata). An overview of the type of swabs and laboratory tests is given in (Table 2-Table 5).

11 Exposure

11.1 Exposure definition

The exposure of interest is influenza vaccination administered during the season 2018-19.For all objectives, the following exposure definitions will be used.

Scenario A:
An individual aged ≥9 years, or a child aged <9 who has been fully vaccinated (at least two injectable doses or one LAIV dose) during the previous influenza season will be considered as

- **vaccinated** with the influenza vaccine of interest if he/she has a record of influenza vaccine administration >14 days before ILI/SARI symptom onset
- **partially vaccinated** if he/she has a record of influenza vaccine administration ≤14 days before ILI/SARI symptom onset
- **unvaccinated** if he/she has no influenza vaccine record for the current season
- **potentially vaccinated** if the positive vaccination status is based on recall alone and cannot be confirmed by registers, or is otherwise ambiguous.

Scenario B:
A child aged < 9 years who has not been fully vaccinated (see above) during the previous influenza season will be considered as

- **vaccinated** with the influenza vaccine of interest if >14 days have elapsed since the second record of injectable vaccination or since the first record of LAIV vaccination during the current season
- **partially vaccinated**
    - after the first record of injectable vaccination until the second record of vaccination during the current season
    - during the first 14 days after the second record of injectable vaccination or the first record of LAIV vaccination during the current season
- **unvaccinated** until the first vaccination record during the season
- **potentially vaccinated** if the positive vaccination status is based on recall alone and cannot be confirmed by registers, or is otherwise ambiguous.

Note 1: The *partially* and *potentially* vaccinated groups will be excluded from primary analysis. The significance of the partially vaccinated subjects will be assessed in sensitivity analyses.
Note 2: If no information on exposure in previous season is available in the dataset, the exposure definition 'scenario A' will be used for all subjects.
Note 3: For cohort studies, vaccination status will be treated as time-varying variable whereas for the case-control studies, vaccination status is a fixed variable.

11.2 Source of exposure information

The sources to obtain information on the exposure status were either vaccine registers, medical records or, vaccination cards. (see Table 2-Table 5). Patients for whom the vaccination status is based on recall only, not verified based on vaccination register, medical record or vaccination card are considered ‘potentially vaccinated’ (see Section 11.1), and will be discarded from analysis (see sections 14.1.1 and 14.2.1).

11.3 Expected influenza vaccine brands

The vaccine types and vaccine brands that are expected to be used in the study areas are summarized in Table 9.

Table 9. Expected vaccine brands and type – all studies, 2018/19

	Site	MUV	CIRI-IT	ISS	RCGP/ UNIS	MUV	HUCH	BIVE-HOSP	NIID	VHUH	FISABIO	THL	UoA	CIRI-IT
	Country	Austria	Italy	Italy	UK	Austria	Finland	Italy	Romania	Spain	Spain	Finland	Greece	Italy
	Study design	TND	TND	TND	TND	TND	TND	TND	TND	TND	TND	Cohort	Cohort	Cohort
	Setting	PC	PC	PC	PC	HOSP	HOSP	HOSP	HOSP	HOSP	HOSP			
	Approved indication													
TIV Brands														
Afluria	5 years and older		X	X				X						X
Agrippal	6 months and older	X	X	X				X		X				X
Influvac	6 months and older	X	X	X	X	X		X	X		X			X
Vaxigrip	6 months and older		X	X				X						X
Intanza	60 years and older													
QIV brands														
Fluarix tetra	6 months and older	X	X	X	X	X		X		X				X
Influvac tetra	3 years and older	X			X	X			X					
Vaxigrip tetra	6 months and older	X	X	X	X	X	X	X	X			X	X	X
aTIV Brands														
Fluad	65 years and older	X	X	X	X	X		X		X	X			X
LAIV brands														
Fluenz tetra	24 months to 17 years	X			X	X						X		

TIV: Trivalent non-adjuvanted; QIV: Quadrivalent inactivated; aTIV: Trivalent adjuvanted; LAIV: Quadrivalent live attenuated

12 Covariates

The additional covariates collected for adjustment are age, sex, presence of at least one chronic condition, pregnancy, number of GP consultations or hospitalizations, and vaccination status in the previous season. An overview of the covariates are given in Table 10 for the TND studies and in Table 11 for the cohort studies.

Table 10. Data collected on covariates – test-negative design studies, 2018/19

Site	MUV	CIRI-IT	ISS	RCGP/ UNIS	MUV	HUCH	BIVE-HOSP	NIID	FISABIO	VHUH
Country	Austria	Italy	Italy	UK	Austria	Finland	Italy	Romania	Spain	Spain
Setting	PC	PC	PC	PC	HO	HO	HO	HO	HO	HO
Age at symptom onset (1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes(2)	Yes
Sex	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Presence of at least one chronic condition	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pregnancy	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Number of hospitalizations in the last 12 months	No	Yes (3)	Yes(3)	Yes	No	Yes	Yes (3)	Yes	Yes	Yes
Number of primary care consultations in the last 12 months	No	No	Yes	Yes	No	Yes	Yes(3)	Yes	Yes (4)	Yes
Receipt of influenza vaccination in 2017-18	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(1) Age in months for children < 1 year, otherwise age in years
(2) Age at hospital admission
(3) Number of hospitalized for any of the chronic conditions of interest (see Annex 3) in the last 12 months
(4) Number of primary care visits during the last 3 months

Table 11. Data collected on covariates – cohort studies, 2018/19

Site	**THL**	**UoA**		**CIRI-IT**
Country	Finland	Greece		Italy
		Pregnant women	Young children	
Age at season onset/cohort inclusion [1]	Yes	Yes	Yes	Yes
Sex	Yes	Yes	Yes	Yes
Presence of at least one chronic condition	Yes	Yes	?	Yes
Pregnancy	n/a	Yes	n/a	Yes
Number of hospitalisations in last 12 months	Yes	No	No	Yes
Number of primary care consultations in the last 12 months	Yes[2]	No	No	No
Receipt of influenza vaccination in 2017-18	Yes	No[3]	n/a	Yes

(1) Age in months for children < 1 year, otherwise age in years

(2) Likely to be an underestimate as private care visits are not counted and follow-up visits are not distinguished from new visits

(3) Any history of influenza vaccination

12.1 Age

Age in years (months for children <1year) at symptom onset.

12.2 Sex

Male or female.

12.3 Date at symptom onset/calendar time

To adjust for time, date at ILI/SARI symptom onset will be used for cohort studies whereas calendar time (in weeks) will be used for cohort studies.

12.4 Chronic conditions

Chronic conditions will be defined as the presence of at least one chronic condition as not all study sites provide information on chronic conditions separately. The chronic conditions include obesity (BMI ≥30) but exclude smoking and pregnancy. The definitions of the chronic conditions are given in Table 12. Listings of chronic conditions included per study site can be found in ANNEX 3.

Table 12. Definitions of chronic conditions

Condition	Definition
Chronic liver disease	Any of the following dg codes (ICD-10)*: B18, K70-74, K75.0-75.1, K75.3-75.9, K76-77 INCLUDING: Alcoholic liver disease, Toxic liver disease, Hepatic failure, Chronic hepatitis (viral & other), Fibrosis and cirrhosis of liver, Other inflammatory liver diseases, Other diseases of liver EXCLUDING: Clinically insignificant liver cysts
Diabetes	Any of the following dg codes (ICD-10)*: E10-E14, O24 INCLUDING: Any form of diabetes, including sequelae & DM in pregnancy
Cardiovascular diseases	Any of the following dg codes (ICD-10)*: A52.0, B37.6, I01-02, I05-09, I11.0, I13.0, I13.2, I20-25, I26-28, I30-43, I44-46, I48, I49.0, I49.5, I50-52, I70-71, Q20-Q28 INCLUDING: all conditions of heart & large vessels that are chronic or likely to have chronic sequelae. Cardiovascular syphilis, endo-, myo- and pericarditis, rheumatic fever, chronic rheumatic heart diseases, congenital malformations, hypertensive (renal) diseases with heart failure, ischaemic heart diseases, diseases of pulmonary circulation, atherosclerosis, cardiomyopathies, most conduction disorders, heart failure, aortic aneurysms & dissecation, other heart diseases and their complications. EXCLUDING: uncomplicated hypertension, previous uncomplicated pulmonary embolism (with no lasting cardiac insufficiency), paroxysmal tachycardias, most cases of premature depolarization.
Cancer	Any of the following dg codes (ICD-10)*: C00-97, D37-48, Z85, Z92.3, Z92.6. INCLUDING: All malignant neoplasms (both solid and haematologic) with potential to metastasize, either in treatment, active followup, or <5 years post curative treatment. EXCLUDING: Benign & in situ neoplasms. Basal cell carcinomas. Any cancer previously treated with curative intent & in complete remission for ≥5 years.
Immuno-deficiency or organ transplant	Any of the following dg codes (ICD-10)*: B20-B24, D80–84, D89, Z94 INCLUDING: HIV infections, immunodeficiencies & organ transplants. or iatrogenic: ≥2 week systemic treatment, in the 3 months preceding symptom onset, with any of the following: corticosteroid (≥20 mg prednisolone daily or equivalent), ciclosporin, tacrolimus, mycophenolate, methotrexate, azathioprine, TNF-α blockers and other biological or cytostatic drugs with immunosuppressive effect EXCLUDING: Disorders of the immune system which do not lead to immunosuppression (e.g. some autoimmune conditions).

Table 12. Definitions of chronic conditions, continued

Condition	Definition
Lung disease	Any of the following dg codes (ICD-10)*: A15-16, A19, A31.0, B33.4, E84.0, J40-47, J60-70, J80-84, J85-86, J90-91, J92.9, J93-94, J95-99 INCLUDING: TB (pulmonary, miliary but not that of other systems), atypical mycobacteria, cystic fibrosis, asthma, COPD, bronchiectasis and other chronic sequelae of infections, chronic lung diseases due to external agents, interstitial lung diseases, pleural diseases, respiratory failure. EXCLUDING: acute respiratory infections, lung cancer, diseases of pulmonary circulation, pleural plaques without asbestos, previous uncomplicated pneumothorax.
Anemia	Any of the following dg codes (ICD-10)*: D50-D64 diagnosed before the onset of symptoms. EXCLUDING: coagulopathies, uncomplicated hypersplenism, hepato/splenomegaly (D65-69, D70-77, D80-84, D86, D89)
Renal disease	Any of the following dg codes: (ICD-10)*: I12-13, M10.30, N00-19, N20.0, N25-27, N28.0, N28.9, Q63.9, Z90.5 EXCLUDING: Clinically nonsignificant kidney cysts
Dementia	Any of the following dg codes (ICD-10)*: F00-03, F05.1, G30-31 EXCLUDING delirium w/o underlying dementia, hydrocephalus.
History of stroke	Any of the following dg codes (ICD-10)*: I61-64, I67.8, I69, G93.1 INCLUDING: both ischaemic and haemorrhaegic strokes and anoxic brain damage. Also counting previous episodes and clear ischaemic findings seen in cranial imaging (even if fully recovered / no symptoms).
Rheumato-logic diseases	Any of the following dg codes (ICD-10)*: M05–09, M13, M30–36, M45 INCLUDING rheumatoid diseases with presumed autoimmune origin and primarily musculoskeletal presentation. EXCLUDING: arthrosis, gout, scoliosis, infectious conditions etc.
Obesity	BMI ≥30 or the dg codes (ICD-10)*: E66, E68 EXCLUDING: local adiposity and "other hyperalimentation" (=vitamin overdoses etc.)

BMI: body mass index; ICD: International classification of diseases

*or corresponding codes in other diagnostic coding systems.

12.5 Pregnancy

Pregnancy (any trimester) at symptom onset: yes versus no.

12.6 Number of hospitalizations

The number of hospitalizations in the previous 12 months will be categorized as “0”, “1 to 2” and “more than 2”. The number of hospitalization is used as a proxy for the severity of chronic conditions.

12.7 Number of primary care consultations

The number of primary care consultations in the previous 12 months (not counting follow-up visits for the same cause) will be categorized as “0”, “1 to 5” and “more than 5”. For FISABIO, only the number of primary care visits in the previous 3 months is available. For FISABIO, this variable will be categorized as “0”, “1 to 2” and “more than 2”. This variable is used as a proxy for health care utilization.

12.8 Vaccination status in previous season

Influenza vaccination status in the previous season will be defined as having received influenza vaccination (any influenza vaccine) during season 2017/2018 as reported in the dataset.

13 Data management

13.1 DRIVE Electronic Study Support Application (ESSA)

The final study data will be uploaded by the DRIVE research study sites to the DRIVE Research server using the DRIVE Electronic Study Support Application (DRIVE ESSA), a password protected web application serving the following purposes:

- Aiding research sites to do the quality assurance of their data by automatically performing data quality checks
- Providing visual summaries of the data
- Allowing research sites to share the visual summaries and tables for monitoring purposes
- Allowing research sites to safely upload their data to the central DRIVE Research Server for statistical analysis

The data flow to the DRIVE Research server is described in Figure 1. The interim and final study data is uploaded by the DRIVE research study sites to the ESSA Server. The DRIVE research study site can decide to share data for monitoring or to transfer the final data to the DRIVE Research Server for statistical analysis. The DRIVE ESSA also aids the research sites providing TND data to do the quality assurance of their data by automatically performing data quality checks and providing visual summaries of their data. The DRIVE ESSA performs 7 different types of quality checks, related to compliance with minimal data requirements, the presence of duplicated records, variable formats and implausible values, inconsistencies between variables and missing values. In addition to the quality checks, the DRIVE ESSA provides seven different data visualizations, summarizing the number of vaccinated subjects over time, the distribution of vaccine brands, the number of cases and controls over time, the age-gender pyramid and the distribution of covariates (sex, age, influenza vaccination previous season, number of hospitalizations during the last 12 months, and presence of at least 1 chronic condition) among cases and controls. More information on the DRIVE ESSA can be found in the DRIVE ESSA user manual (ANNEX 2). Similar data quality checks will be performed for the cohort studies. Performing quality checks for the cohort studies is currently not yet implemented in the DRIVE ESSA and will be done by writing separate data management scripts.

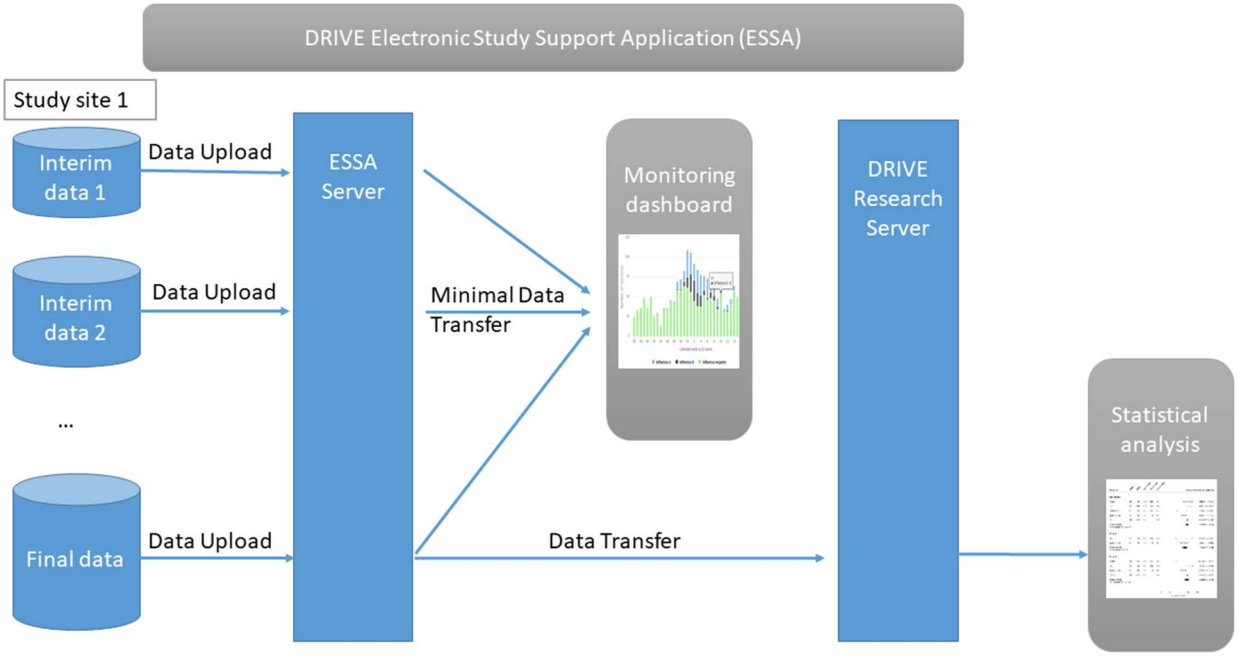


Figure 1. DRIVE Electronic Study Support Application: data flow

13.2 DRIVE Research server

The DRIVE Research server, provided by P95, is a highly secured IT environment and network with strict rules for data access. The architecture of the DRIVE research server is given in Figure 2. The general architecture of the DRIVE research server has three compartments: the data import compartment, the data analysis compartment and the evidence export compartment. The DRIVE research server is only accessible through the secure file transfer protocol (with upload capability to the data import compartment and download capability out of the data export compartment) and the remote desktop protocol allowing data analysts/statisticians to log into the data analysis compartment. The transfer of any data between the different compartments is done solely by the server administrator where data privacy assessments should be carried out if deemed necessary. Every interaction on the DRIVE research server will be logged, and these logs will be accessible upon request.

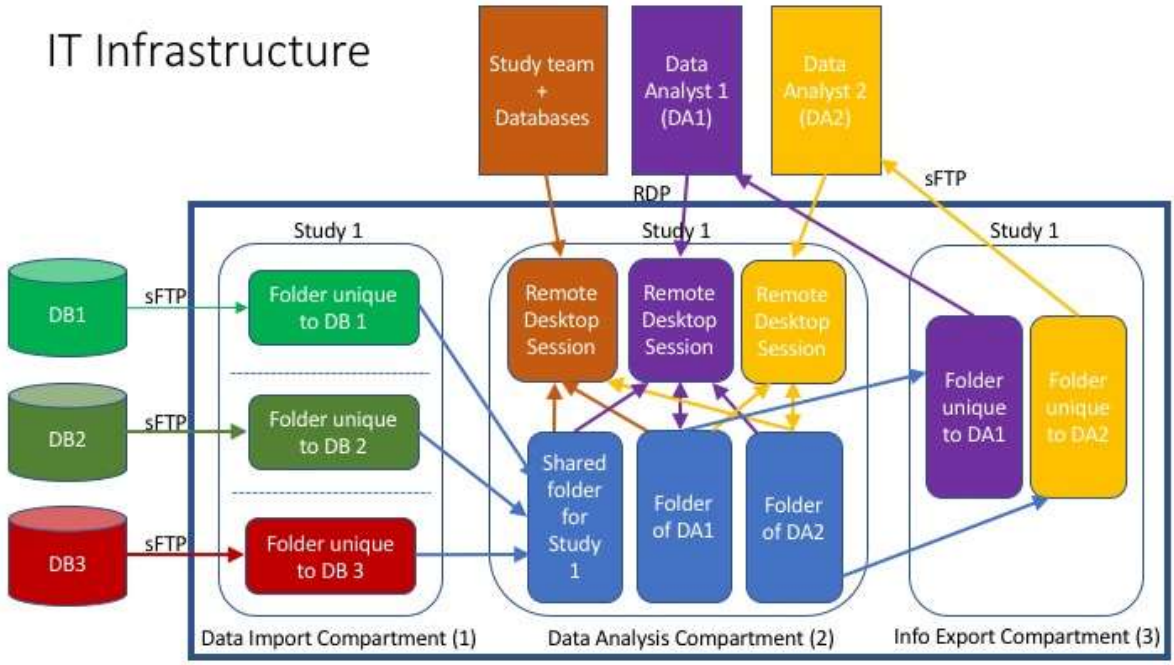
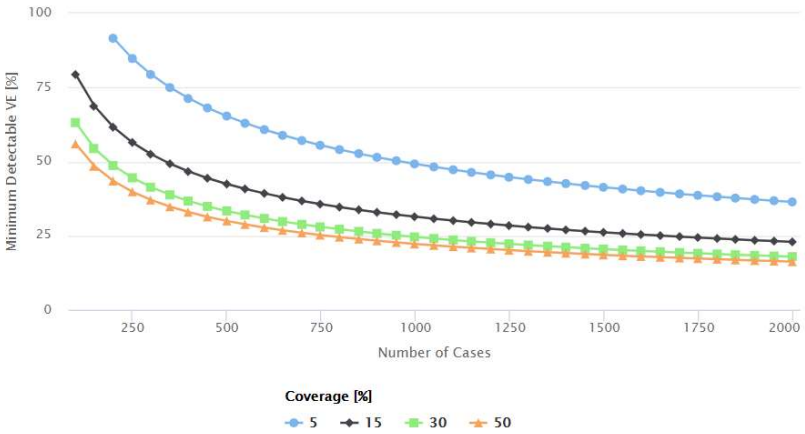


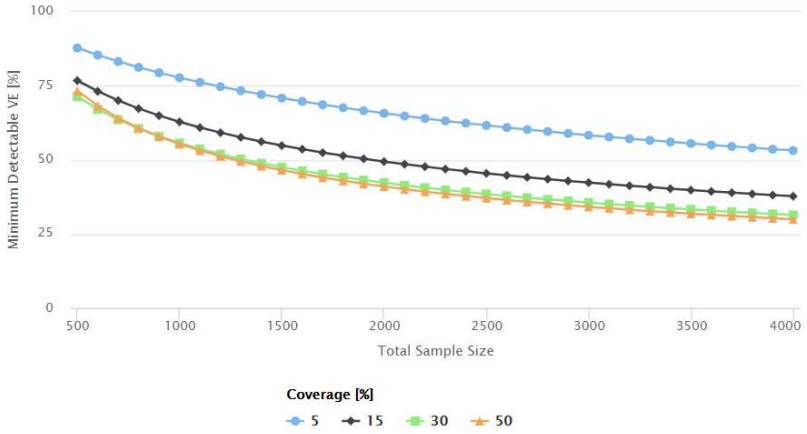
Figure 2. DRIVE Research server: architecture

14 Sample size considerations

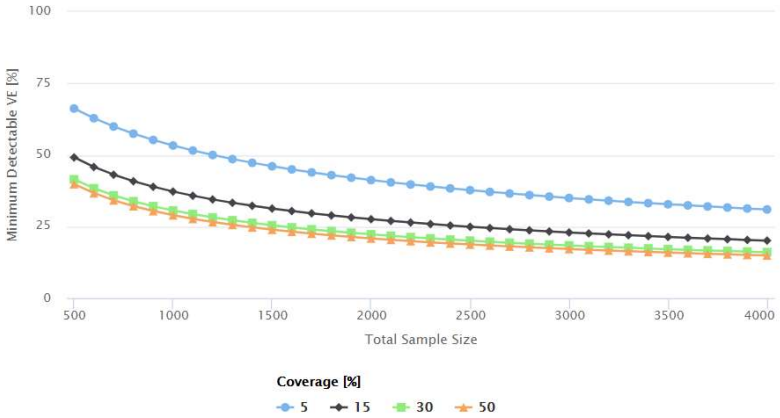
The minimal detectable overall VE, or the smallest VE that can be detected as significantly greater than zero, for a range of samples sizes for TND and cohort designs is given in Figure 3 and Tables Table 13-Table 14. The calculations are performed assuming 80% power, two-sided 95% confidence levels and overall vaccination coverages of 5%, 15%, 30% and 50%. For TND, it is additionally assumed to have a 1:1 control per case allocation ratio. For the cohort studies, it is additionally assumed to have attack rates among the unvaccinated of 7% (reflective of the attack rate in adults) and 25% (reflective of the attack rate in children).



Test-negative design



Cohort design (attack rate = 7%)



Cohort design (attack rate = 25%)

Figure 3. Minimal detectable overall Vaccine effectiveness (VE) for test-negative and cohort design studies, assuming 80% power, two-sided 95% confidence intervals and overall vaccination coverage of 5%, 15%, 30% and 50%. For the test-negative design, a 1:1 control per case allocation ratio is assumed. For the cohort design, attack rates of 7% and 25% are assumed.

Table 13. Minimum detectable overall Vaccine effectiveness (VE) for the test-negative design studies, assuming 80% power, two-sided 95% confidence intervals, a 1:1 control per case allocation ratio and overall vaccination coverage of 5%, 15%, 30% and 50%.

	Minimum detectable VE			
Number of cases	**5% Coverage**	**15% Coverage**	**30% Coverage**	**50% Coverage**
100	NA	79.35	63.16	55.98
200	91.57	61.58	48.67	43.44
500	65.35	42.45	33.35	30
750	55.49	35.64	27.95	25.22
1000	49.21	31.39	24.59	22.23
1250	44.72	28.4	22.24	20.13
1500	41.32	26.15	20.47	18.54
2000	36.4	22.92	17.93	16.26

Table 14. Minimum detectable overall Vaccine effectiveness (VE) for cohort design studies, assuming 80% power, two-sided 95% confidence intervals, 7% and 25% attack rate in the unvaccinated and overall vaccination coverage of 5%, 15%, 30% and 50%.

		Minimum detectable VE			
Sample size	**Attack rate %**	**5% Coverage**	**15% Coverage**	**30% Coverage**	**50% Coverage**
500	7	87.66	76.69	71.29	73.16
1000	7	77.65	62.76	55.64	55.29
1500	7	70.78	54.77	47.47	46.48
2000	7	65.64	49.38	42.2	40.96
2500	7	61.58	45.4	38.43	37.07
3000	7	58.26	42.29	35.54	34.13
3500	7	55.48	39.78	33.24	31.81
4000	7	53.09	37.69	31.35	29.92
500	25	66.2	49.12	41.46	39.8
1000	25	53.28	37.21	30.6	28.95
1500	25	46.04	31.29	25.45	23.91
2000	25	41.21	27.56	22.27	20.85
2500	25	37.69	24.93	20.06	18.74
3000	25	34.96	22.94	18.4	17.16
3500	25	32.77	21.38	17.11	15.93
4000	25	30.96	20.1	16.05	14.93

We recommend a minimum of 200 influenza positive cases for TND studies and a minimum of 1000 subjects for cohort studies. As data from different sites will be pooled and as capacity building is an ongoing activity within DRIVE, smaller sample sizes per site are allowed. To not spread resources too thinly, it is recommended to select study sites that are expected to provide at least 100 cases in the case of TND studies and 500 subjects in the case of cohort studies. A user-friendly web-application to perform sample size calculations for IVE studies has been developed and is freely available from http://apps.p-95.com/drivesamplesize/.

15 Statistical analysis

15.1 Site-specific analysis: test-negative design studies

15.1.1 Attrition diagram

Records will be discarded from analysis when:

- Date of ILI/SARI onset is outside the study period (see Section 7)
- Subjects do not adhere to the study in- and exclusion criteria (see Section 8).
- The ILI/SARI episode is not the first episode from recurrent episodes within the study period
- Subjects have missing information on the outcome of interest or exposure or any of the covariates for adjustment (see Section 11)
- Subjects that are potentially vaccinated (see Section 11.1)

Subjects that are partially vaccinated (see Section 11.1) will be excluded from the primary analysis. The impact of partially vaccinated subjects will be assessed in sensitivity analysis.

When a covariate contains a large percentage of missing data (>= 10%), no adjustment will be made for that covariate to avoid losing too many records, unless the covariate adjustment is considered more important than the information loss. In that case or when the covariate is not available, missing information on that specific covariate will not be a reason for exclusion.

For every TND study site (n = 11), an attrition diagram as outlined in Figure 4 will be created. The attrition diagram describes the amount of records excluded from the statistical analysis by reason of exclusion.

It is expected that some study sites will provide data to DRIVE with the exclusion criteria applied at study recruitment or during data cleaning before sharing data while for other sites the exclusion criteria will be applied by the DRIVE study team.

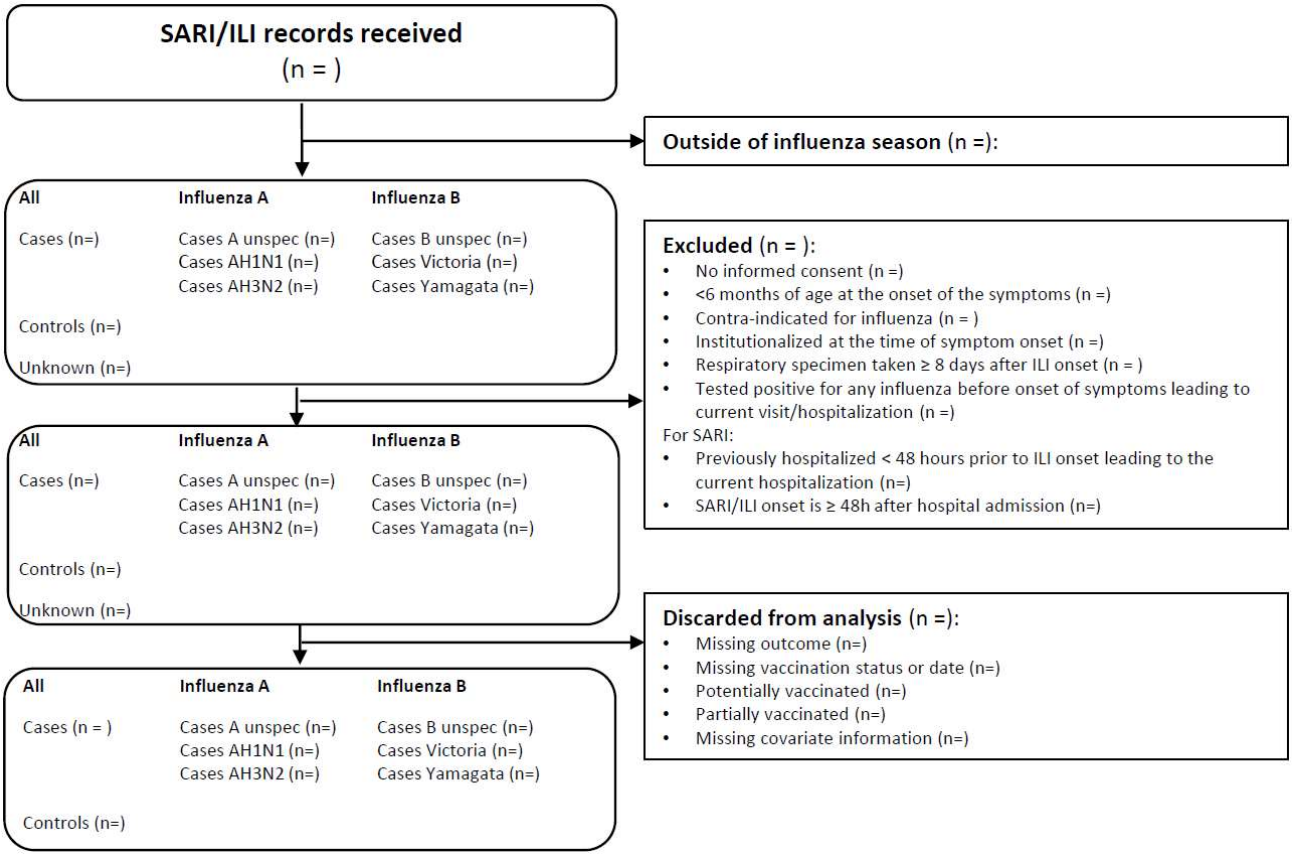


Figure 4. Attrition diagram, study site A

15.1.2 Descriptive analysis

For every TND study site, visualizations based on the final data (e.g. data for analysis) will be created similar as the ones generated by the DRIVE ESSA (ANNEX 2), including:

- Pie chart of the distribution of vaccine brands
- Cumulative number of vaccinated subjects over time
- Number of controls and laboratory-confirmed influenza infections (by type and by subtype/lineage) over time
- Distribution of covariates among cases and controls

For every TND study site (n = 11), a table based on the final data will be created with characteristics of cases and controls as outlined in Table 15.

Table 15. Table shell: characteristics of laboratory confirmed cases, overall and by influenza type, and controls, study site A (setting), 2018/19

Characteristic	Cases N (%) all	A (H1N1/H3N2)	B (Vict/Yama)	Controls N (%)
Age group				
6m-17 yr				
18-64 yr				
≥65 yr				
Sex				
female				
male				
At least 1 chronic condition				
Yes				
Cardiovascular disease				
Lung disease				
Diabetes				
Immunodeficiency or organ transplant				
Chronic liver disease				
Cancer				
Anemia				
Renal disease				
Dementia				
Stroke				
Rheumatologic diseases				
Obesity				
No				
Unknown				
Pregnancy				
Yes				
No				
Unknown				
Number of primary care visits in the previous 12 months				
0				
1-5				
>5				
Unknown				
Number of hospitalizations in the previous 12 months				
0				
1-5				
>5				
Unknown				
Influenza vaccination status in previous season				
Vaccinated				
Unvaccinated				
Unknown				
Influenza vaccination status in current season				
Vaccinated				
Afluria				
Agrippal				
…				
undefined				
Unvaccinated				
Total				

15.1.3 Influenza vaccine effectiveness estimation

For every TND study site,crude and confounder-adjusted IVE (any influenza vaccine, by brand and by vaccine type) against laboratory-confirmed influenza (any, by influenza type and subtype/lineage) will be estimated stratified by age (6m-17yr, 18-64 yr, >=65yr), as

$$VE = (1 - OR) \times 100\%,$$

where *OR* denotes the confounder-adjusted odds ratio, comparing the odds of vaccination among influenza-positive study participants to the odds of vaccination among influenza-negative study participants.

Confounder-adjusted IVE estimates will be derived from multivariable logistic regression models. A fixed set of confounders will be considered for each individual site, including sex, a smooth function of age, a smooth function of symptom onset date, presence of at least one chronic condition, pregnancy, number of primary care visits (FISABIO: “0”, “1 to 2” and “2 or more”; all other sites: “0”, “1 to 5” and “5 or more”) in the previous 12 months (for primary care studies) or number of hospitalizations (“0”, “1 to 2” and “2 or more”) in the previous 12 months (for hospital based studies) and influenza vaccination in the previous season. This set of confounders is available for the majority of study sites. (See also Table 10 for the confounders available per site).

The analysis to estimate brand-specific IVE will account for the differences in approved indications (see Table 6), discarding from the analysis subjects for which the vaccine brand of interest is not indicated.

The analysis will be a complete case analysis, dropping records with missing information for the outcome, exposure of interest or any of the covariates. The smooth functions of age and symptom onset date will be modelled by penalized cubic regression splines and estimated using restricted maximum likelihood for smoothness selection [6].

For sites for which some confounders are entirely missing, the IVE estimates will be confounder-adjusted to the extent possible.

For the trivalent vaccines and trivalent vaccine types (i.e. trivalent non-adjuvanted, trivalent adjuvanted, trivalent high-dose), an additional IVE estimate against any vaccine subtype/lineage included in the vaccine will be obtained.

For each study site, the estimates will be presented as in Table 16. A similar table as Table 16 will be made for the crude IVE estimates. This will yield a total of 22 tables for the TND studies.

Table 16. Table shell: confounder-adjusted influenza vaccine effectiveness [95% confidence intervals], study site A (setting), 2018/19

Study site (setting)	Influenza Vaccine Effectiveness [95% CI]*						
	Any	**Any vaccine strain**	**AH1N1**	**AH3N2**	**B**	**B Victoria**	**B Yamagata**
Age group							
6m-17 yr							
Any vaccine							
Vaccine brand							
Afluria(1)							
Agrippal							
…							
Vaccine type							
Trivalent adj							
Trivalent non-adj							
…							
18-64 yr							
Same as above							
…							
>=65 yr							
Same as above							
…							

CI: confidence interval
(1) all brands available at the study site

15.1.4 Sensitivity analysis

The following sensitivity analysis will be conducted:

Partially vaccinated subjects:

- partially vaccinated subjects (see Section 11.1) will be considered unvaccinated.
- partially vaccinated subjects (see Section 11.1) will be considered vaccinated

Time between ILI/SARI onset and swab:

- subjects will be excluded when the respiratory specimen was taken ≥ 4 days after ILI/SARI onset

15.2 Site-specific analysis: cohort studies

15.2.1 Attrition diagram

Records will be discarded from the analysis when:

- Date of ILI/SARI onset is outside the study period (see Section 6)*
- Subjects do not adhere to the study in- and exclusion criteria (see Section 9).
- Subjects have missing information on the outcome of interest, exposure or any of the covariates for adjustment (see Section 11)
- Subjects that are potentially vaccinated (see Section 11.1)

Note 1: Date of ILI/SARI onset cannot be verified in the THL register-based cohort study. Only the date of influenza test is available.

Note2: The in- and exclusion criteria are different for the different cohort studies.

When a covariate contains a large percentage of missing data (>= 10%), no adjustment will be made for that covariate to avoid losing too many records, unless the covariate adjustment is considered more important than the information loss.

For every cohort study, an attrition diagram similar to the one given in Figure 4 will be created. The attrition diagram describes the amount of records excluded from the statistical analysis by reason of exclusion.

It is expected that some study sites will provide data to DRIVE with the exclusion criteria applied before sharing data while for other sites the exclusion criteria will be applied by the DRIVE study team.

15.2.2 Descriptive analysis

For every cohort study, visualizations based on the final data (e.g. data for analysis) will be created similar as the ones created for the TND studies (see Section 14.1.2), including:

- Pie chart of the distribution of vaccine brands
- Cumulative number of vaccinated subjects over time
- Number of laboratory-confirmed influenza infections (by type and subtype/lineage) over time
- Distribution of covariates among exposed and unexposed subjects

For every cohort study, a table based on the final data will be created with characteristics of the exposed and unexposed subjects as outlined in Table 17. Table 17 uses the UoA pregnancy cohort (Greece) as an example. Similar tables will be made for the other cohort studies.

Table 17. Table shell: characteristics of the exposed and unexposed subjects, 2018/19 . UoA (Greece, pregnancy cohort) as example.

Characteristic	vaccinated		Unvaccinated	
	N	Person time	N	Person time
Age group				
18-29 yr				
30-45 yr				
≥65yr				
Ethnicity				
Greek				
Roma				
Immigrant				
At least 1 chronic condition				
Yes				
No				
Unknown				
Influenza vaccination in previous season				
Vaccinated				
Unvaccinated				
Unknown				
Number of children < 5 yr				
0				
1-2				
>2				
Unknown				
Influenza vaccination status in current season				
Vaccinated				
Afluria				
Agrippal				
…				
undefined				
Partially vaccinated				
Unvaccinated				
Total				

15.2.3 Influenza vaccine effectiveness estimation

For every cohort study, crude and confounder-adjusted IVE (any influenza vaccine, by brand and by vaccine type) against laboratory-confirmed influenza (any, by influenza type and subtype/lineage) will be estimated stratified by age (6m-17yr, 18-64 yr, >=65yr), as

$$VE = (1 - RR) \times 100\%,$$

where *RR* denotes the confounder-adjusted relative risk, comparing the influenza incidence among the vaccinated subjects to the influenza incidence among the unvaccinated subjects.

Confounders include sex, a smooth function of age, a smooth function of calendar week, presence of at least one chronic condition, pregnancy, number of primary care visits (“0”, “1 to 5” and “5 or more”) in the previous 12 months (for primary care studies) or number of hospitalizations (“0”, “1 to 2” and “2 or more”) in the previous 12 months (for hospital based studies) and influenza vaccination in the previous season, whenever available (See Table 11).

The analysis will be a complete case analysis, dropping records with missing information for the outcome, exposure of interest or any of the covariates. The smooth functions of age and calendar time will be modelled by penalized cubic regression splines [7] estimated using restricted maximum likelihood for smoothness selection [6]

The crude and confounder-adjusted estimates will be presented as in Table 16. For the cohort studies, a total of 6 tables will be generated.

15.2.4 Sensitivity analysis

The following sensitivity analysis will be conducted:

a) Partially vaccinated subjects:

- partially vaccinated subjects (see Section 11.1) will be considered unvaccinated.
- partially vaccinated subjects (see Section 11.1) will be considered vaccinated

b) Time between ILI/SARI onset and swab:

- subjects will be excluded when the respiratory specimen was taken ≥ 4 days after ILI/SARI onset

For the register-based cohort study (THL, Finland), sensitivity analysis b) will not be considered as the information on ILI/SARI onset is missing.

15.3 Pooled analysis

15.3.1 Inclusion of influenza vaccine effectiveness estimates

Only estimates provided by the TND studies will be considered for obtaining pooled estimates stratified by age group (6m-17yr, 18-64 yr, >=65yr) and setting (primary care, hospital).

The clinical cohort studies will not be considered for inclusion in the pooled analyses as they concern different populations compared to the general population covered by the TND studies (UoA Greece; pregnant women and their infants; CIRI-IT Italy; healthcare workers).

The population-based cohort study (THL, Finland) will also not be considered for inclusion in the pooled analysis for this year as primary care based or hospitalized laboratory-confirmed influenza cases cannot be disentangled.

15.3.2 Pooled data: descriptive analysis

For the TND data, tables based on the pooled data will be created with characteristics of cases and controls similar to Table 15, stratified by healthcare setting (primary care, hospital). These two tables on the pooled data will additionally contain information on the distribution of cases and controls across the different study sites.

Additional tables will be created describing the characteristics of the subjects by exposure as outlined in Table 18 to obtain insight into potential bias by indication. As different brands have different approved indications (see Table 9), the unexposed group that will serve as a basis for comparison might be slightly different. A separate table will be created by healthcare setting and age group.

Table 18. Table shell: characteristics of the exposed and unexposed subjects by brand, setting x age group, 2018/19

	Brand 1		Brand 2 ….	
	exposed	**unexposed**	**exposed**	**unexposed**
Age at symptom onset				
Average (SD)				
Median				
Sex				
female				
male				
At least 1 chronic condition				
Yes				
Cardiovascular disease				
Lung disease				
Diabetes				
Immunodeficiency or organ transplant				
Chronic liver disease				
Cancer				
Anemia				
Renal disease				

Dementia
Stroke
Rheumatologic diseases
Obesity
No
Unknown

Pregnancy
Yes
No
Unknown

Number of primary care visits in the previous 12 months
0
1-5
>5
Unknown

Number of hospitalizations in the previous 12 months
0
1-5
>5
Unknown

Influenza vaccination status in previous season
Vaccinated
Unvaccinated
Unknown

Study site
Site 1
Site 2
…

Total

15.3.3 Meta-analysis

Random effects meta-analysis (RE MA) [8] will be used to pool the site-specific confounder-adjusted IVE estimates as given in Table 16. Pooled estimates will be stratified by age group (6m-17yr, 18-64 yr, >=65yr) and setting (primary care, hospital). Random effects meta-analysis will be performed on the log-transformed odds ratio (OR) estimates. Restricted maximum likelihood (REML) will be used to obtain the pooled (meta-analyzed) estimate (and 95% confidence intervals), as the REML estimator outperforms other RE MA estimators in terms of bias and statistical efficiency [9]. The estimates (and 95% confidence intervals) will then be back-transformed to obtain the pooled IVE estimate (and 95% confidence intervals), expressed in %.

15.3.4 Quantifying between-study heterogeneity

An indication for the heterogeneity among estimates from different study sites will be obtained by calculating I^2 according to Higgins et al [10]. The I^2 statistic is to be interpreted as the proportion of total variation in the estimates of treatment effect that is due to heterogeneity between studies. This measure will be used as a summary measure of the between-study heterogeneity and not to decide on the appropriateness of pooling as the RE MA model accounts for different levels of between-study heterogeneity.

15.3.5 Outlier and influence analysis, and exploring reasons for potential outlying studies

For every meta-analysis performed, the potential impact of outliers and influential estimates on the pooled estimate will be evaluated. Studentized deleted residuals *r* will be used to identify outliers in the meta-analysis. Site-specific IVE estimates will be considered outlying from meta-analysis when $|r| > 2.5$, where $|r|$ indicates the absolute value of the residual.

The standardized DFBETAs statistic will be used to identify influential estimates, examining the change in the averaged IVE from the random-effects model when excluding one site-specific estimate in turn. Site-specific estimates will be considered influential from meta-analysis when |DFBETAs| > $2/\sqrt{n}$, where |DFBETAs| indicates the absolute value of the DFBETAs statistics and *n* is the number of effect estimates [11].

Site-specific estimates that are outlying and influential, will be excluded from meta-analysis and the reason for being outlying will be investigated and documented. The information that will be collected by the DRIVE Quality Control and Audit Committee (QCAC) will be used to evaluate potential reasons for outlying results.

15.3.6 Sensitivity analysis

The following sensitivity analysis will be conducted:

a) Partially vaccinated subjects:

- partially vaccinated subjects (see Section 11.1) will be considered unvaccinated.
- partially vaccinated subjects (see Section 11.1) will be considered vaccinated

b) Time between ILI/SARI onset and swab:

- subjects will be excluded when the respiratory specimen was taken ≥ 4 days after ILI/SARI onset

c) Outlying/influential studies:

- Outlying/influential studies will be included in the meta-analysis, if any

16 Presentation of results

For every exposure of interest (any vaccine, by brand, by vaccine type), a separate multi-panel plot will be created, displaying the IVE estimates (any influenza, by type and by subtype/lineage) stratified by age group and setting/design (primary care TND, hospital-based TND, cohort studies). The pooled estimates (from the TND studies) will be represented by diamonds whereas sites-specificl estimates (from the cohort studies) will be represented using error bars. For the pooled estimates, the I^2 statistic will be given as well. Wide confidence intervals (i.e. a width > 40%) will be coloured differently compared to narrow confidence intervals (i.e. width <= 40%) to emphasize that estimates with wide confidence intervals are not considered robust. An example of such a multi-panel plot is given in Figure 5. The plots also make explicit which estimates are still missing for the current season. We anticipate to produce 16 multi-panel plots (any vaccine + 10 brands + 4 vaccine types).

For every pooled estimate, an additional forest plot with the site-specific estimates will be created to support the interpretation of the pooled estimates. An example of such a forest plot displaying different IVE estimates (by type and by subtype/lineage) is given in Figure 6.

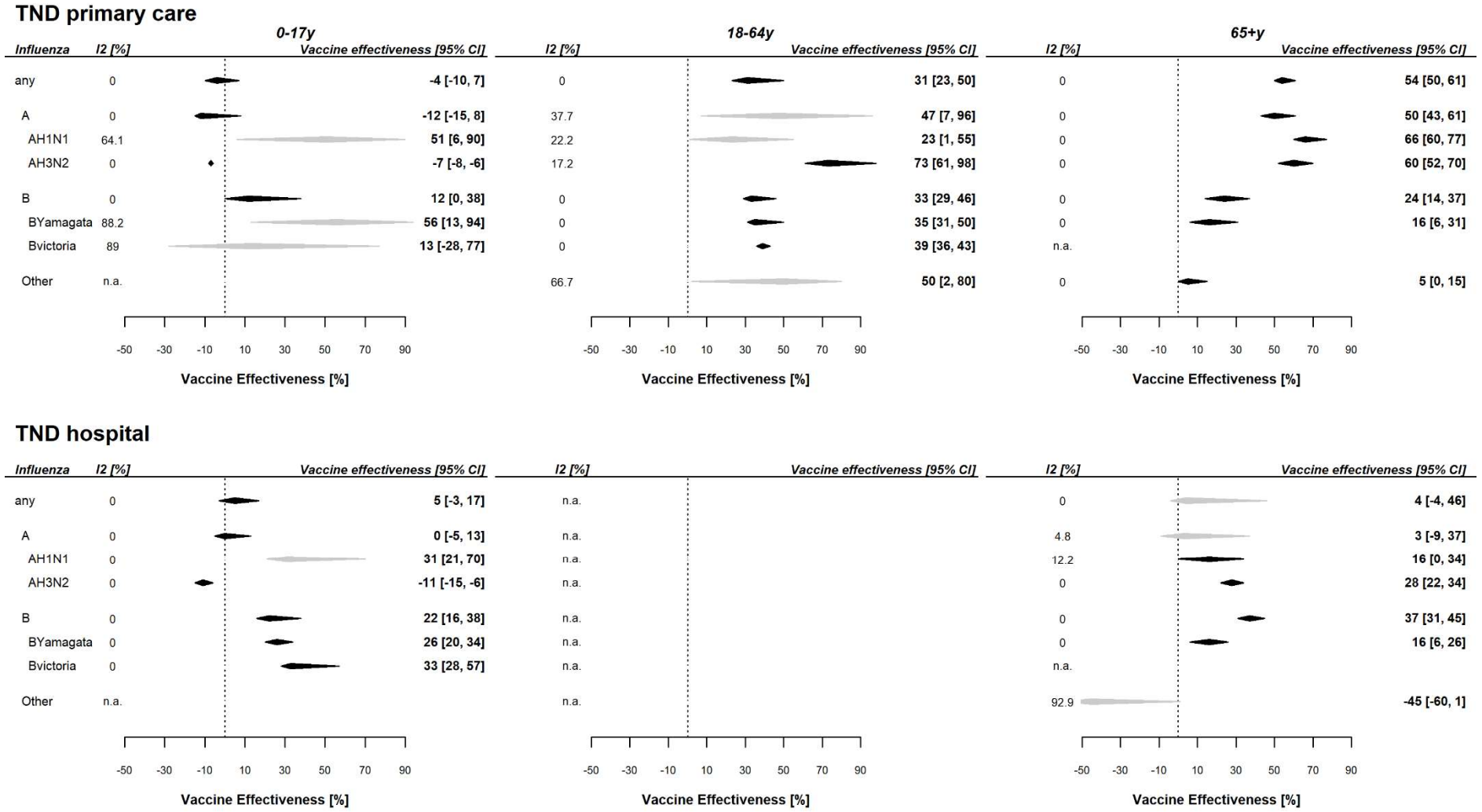


Figure 5. Example of a multi-panel plot, by exposure (artificial example)

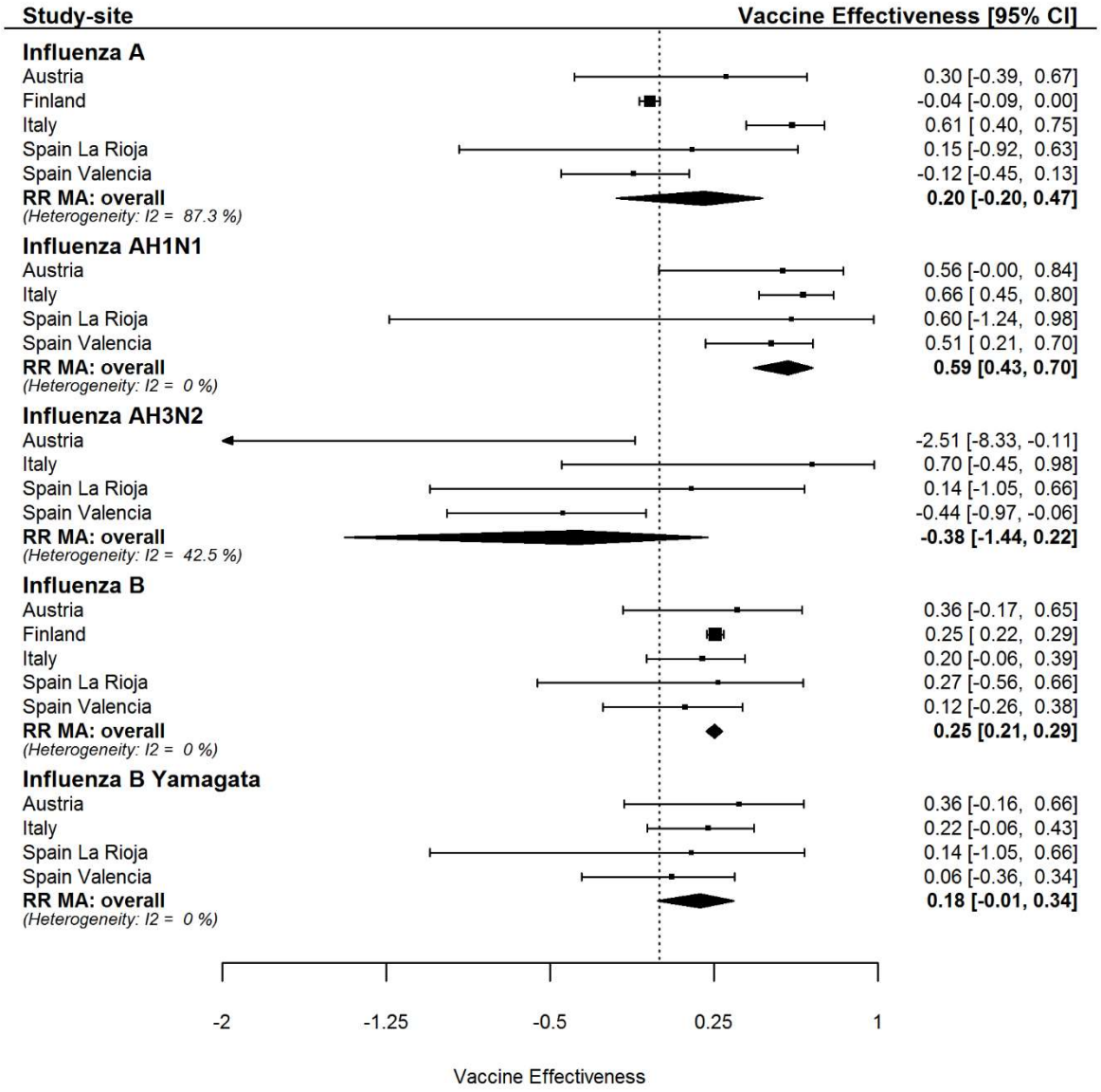


Figure 6. Example of a forest plot associated with pooled Influenza vaccine effectiveness estimates (setting A , age group B , exposure CSee) by influenza type and subtype/lineage (based on data from pilot study)

17 Software

All data management and statistical analyses will be conducted in R 3.5.2.

18 Limitations

The populations of the clinical cohort studies are different from the general population as studies by the TND and Finnish population-based cohort study. As such, the results from the clinical cohort studies cannot be pooled with the results from the other studies. At this moment, it is still unclear how much such clinical cohort studies can contribute to DRIVE's primary objective of estimating brand-specific IVE in Europe.

For some study sites, no information is available on influenza subtypes/lineages, the information on covariates is limited or primary care or hospital based cases cannot be distinguished. It remains to be decided what information is minimally required for obtaining robust IVE estimates, and hence which are the minimum study requirements for the DRIVE studies.

All TND studies closely follow the generic TND study protocol. However, the study sites are still different in many important aspects, including the sampling strategy and covariates available for adjustment. Several potential confounders are currently not available such as socio-economic status and smoking. The covariates 'at least 1 chronic condition' or 'influenza vaccination in the last season' might not be sufficiently granular to allow for proper confounder adjustment.

It is difficult to know the sample size required for brand-specific IVE as it depends on many unknown factors, including the influenza attack rates, vaccination coverage, distribution of brands and (for the pooled estimates) the between-study heterogeneity. For the 2018/19 season, we will likely have sufficient sample size to obtain robust and precise (with CI width < 40%) IVE estimates for some brands, but not for all brands used in Europe in the 2018/19 season. Obtaining sufficient sample for brand-specific IVE estimates will remain a challenge and a careful selection of DRIVE study sites will be required.

Bias by indication is a challenge in IVE studies and will also likely affect the IVE estimates of this season, despite attempts to correct for bias by indication through covariate adjustment. It will be particularly important to understand the target groups for influenza vaccination as well as the target groups for vaccination with specific influenza vaccine brands.

19 Quality control procedures

19.1 Documentation

The following study documents will be generated and are available upon request: Generic site-specific protocols, description of minimum data requirements, season-specific protocol per study site, season-specific SAP, study summary sheets and codebook of variables within the analytical datasets.

19.2 Record retention

Documents that permit evaluation of the conduct of a study and the quality of the data will be retained for a period of 5 years in accordance with Good Participatory Practice (GPP) guidelines. These documents will be retained for a longer period, however, if required by the applicable regulatory requirements or by an agreement between study partners.

The analytical datasets and data analysis scripts used to produce the site-specific IVE estimates will be retained at the participating study sites and made available for quality control upon request. The syntaxes used for the pooled analysis will be documented and made available for quality control upon request.

19.3 Data analysis and results

Quality control of R programs will include a review of the whole process of the results generation:

- Review of all analysis R programs;
- Review of R logs for errors, uninitialized variables and warnings;
- Review of all tables, listings and figures for completeness and correctness.

19.4 Monitoring of quality

The Quality Control and Audit Committee (QCAC) of DRIVE is composed of external quality control advisors (who may or may not represent consortium members). The QCAC will perform a 3-step assessment of the quality of the studies:

- Study conduct: whether study was conducted in compliance with regulatory standards, the site protocol and the local SOPs
- Quality of the data: whether data collected from the field were processed in compliance with the DRIVE Data Management Plan (DMP)

- Quality of the analysis: whether the pooled statistical analysis report matches with the Statistical Analysis Plan (SAP).

To evaluate these points, the QCAC will develop three checklists in agreement with DRIVE WP3 and P95. Based on the evaluation, the QCAC will provide recommendations to DRIVE Steering Committee. The conclusion of QCAC will be described in a quality report and attached to the final study report.

20 Ethics considerations

20.1 Ethics approval

Participating sites obtained ethics committee approval as required. The ethics committee that approved the study at each site and the date of approval are listed in Table 19. For ISS (Italy), separate ethics committee approval was not required as these studies are part of their respective National Influenza Surveillance activities.

Table 199. DRIVE 2018/19 study sites: ethics committees and date of approval

Site	Country	Ethics committee	Date of approval
MUV	Austria	Ethics committee of the MUV	May 4, 2018
HUCH	Finland	Regional Ethics Committee of the Expert Responsibility Area of Helsinki University Hospital	Nov 14, 2018
THL	Finland	Institutional review board of the National Institute for Health and Welfare, Finland	June 2, 2016
UoA	Greece	Ethics Committee of the "Alexandra" General Hospital of Athens	Oct 16, 2018
BIVE-HOSP	Italy	Ethics committee of the Bambino Gesù Children's Hospital, Rome Ethics committee of the Sant'Andrea Hospital, Rome Ethics committee of the University Hospital, Bari Ethics committee of the San Martino Hospital, Genova Ethics committee of the Le Scotte Hospital, Siena	Sept 2018 (all committees)
CIRI-IT (TND)	Italy	Ethics committee of the Liguria Region	Oct 1, 2018
CIRI-IT (HCW)	Italy	Ethics committee of the Liguria Region	Oct 1, 2018
ISS	Italy	Not required, but submitted to ISS Ethics committee for information	Nov 23, 2018
NIIS	Romania	Bioethics committee of the NIIS	Nov 12, 2018
FISABIO	Spain	National Ethics Committee	Dec 21, 2009
VHUH	Spain	Comité Ético de Investigación Clínica del Hospital Universitari Vall d'Hebron	Dec 13, 2018
RCGP/UNIS	UK	NRES Committee West Midlands, Solihull. IRAS project ID: 252081, REC reference: 19/WM/0015	Feb 4, 2019

20.2 Informed consent

At all sites except VHUH and THL informed consent was required. For the THL register-based cohort study, informed consent was not required as the study makes use of secondary data from routine databases.
For the VHUH study, informed consent was not required as no interventions that fall outside the usual practice at VHUH during the influenza season were needed.

21 References

(1) Committee for Medicinal Products for Human Use. Guideline on Influenza Vaccines - Non-clinical and Clinical Module. EMA/CHMP/BWP/310834/2012. In. London: Eur Med Agency, 2016.
(2) **DRIVE consortium**. D7.4 Setting up brand-specific influenza vaccine effectiveness studies in Europe – results of the pilot season 2017/18. Accessible from: https://www.drive-eu.org/wp-content/uploads/2018/12/D7_4_Report-pilot-season-201718_v1.0.pdf; October 2018 October 2018.
(3) ECDC. EU case definitions / Influenza including Influenza A(H1N1). In. Stockholm 2018. Accessible: https://ecdc.europa.eu/en/surveillance-and-disease-data/eu-case-definitions.
(4) **Dhiman N, et al.** Effectiveness of patient-collected swabs for influenza testing. *Mayo Clin Proc* 2012; **87**(6): 548-554.
(5) **Hayward AC, et al.** Comparative community burden and severity of seasonal and pandemic influenza: results of the Flu Watch cohort study. *Lancet Respir Med* 2014; **2**(6): 445-454.
(6) **Marra G, Wood SN.** Practical variable selection for generalized additive models. *Computational Statistics and Data Analysis* 2011: 15.
(7) **Wood S**. *Generalized additive models: an introduction with R*. London: CRC Press, 2017.
(8) **DerSimonian R, Laird N.** Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**(3): 177-188.
(9) **Viechtbauer W.** Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics* 2005; **30**.
(10) **Higgins JP, Thompson SG.** Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**(11): 1539-1558.
(11) **Viechtbauer W, Cheung MW.** Outlier and influence diagnostics for meta-analysis. *Res Synth Methods* 2010; **1**(2): 112-125.

Brand-specific influenza vaccine effectiveness in Europe Statistical Analysis Plan Season 2018-2019

APPENDICES

777363 - DRIVE

Development of robust and innovative vaccine effectiveness

WP7 - IVE studies

TABLE OF CONTENTS

1 Annex 1. Minimum data requirements

Table S.1.1: DRIVE – Minimum dataset for pooled data analysis (case-control studies) v07 14 Feb 2019

Variable	Obligatory	Description	Additional info	Format	Values/coding	Example
idcountry	Obligatory	Country code defined in ISO 3166-1 alpha-2		2 letters text		UK
idstudy	Obligatory	Name of the study		Text		JorviTND
region	Optional	Region name		Text		Wales
idunit	Obligatory (for studies which include >1 GP offices / hospitals)	Identifier of the GP practice or hospital where the patient was seen		Text		JS123
setting	Optional	Type of unit (outpatient, e.g. GP practice, or inpatient, e.g. hospital)		Numeric (Categorical)	1=Outpatient 2=Inpatient 9999=No information	2
id	Obligatory	Patient identification number		Unique integer		101
sex	Obligatory	Sex		Numeric (Binary)	0=Female 1=Male	0
age	Obligatory	Age in years (at the onset of the symptoms)		Numeric		84

agemonths	Obligatory for children <1 year of age	Age in months (only for children <1 years old. Else should not be provided.)		Numeric		6
onsetdate	Obligatory	Date of symptoms onset		dd/mm/yyyy	Date within the study period	29/12/2017
swabdate	Obligatory	Date of swabbing		dd/mm/yyyy	Date within the study period	30/12/2017
visitdate	Obligatory	Date of visit to the GP or admission to the hospital	In hospital, the first point of contact (often, arrival at the emergency room)	dd/mm/yyyy	Date within the study period	30/12/2017
death	Optional	Has the patient died?	During hospitalization or within 30 days after discharge	Numeric (Binary)	0=Alive 1=Dead	0
deathdate	Optional	Date of death		dd/mm/yyyy	Date within the study period	99/99/9999
fever	Optional	Fever or feverishness	A measured fever of ≥38°C or temperature 37-38°C with patient-reported feverishness	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
headache	Optional	Headache		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
myalgia	Optional	Myalgia		Numeric (Categorical)	0=No 1=Yes 9999=No information	0
malaise	Optional	Fatigue/Malaise		Numeric (Categorical	0=No 1=Yes 9999=No information	1

suddenonset	Optional	Sudden onset of symptoms	Within 7 days before admission	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
cough	Optional	Cough		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
diffbreath	Optional	Difficulty breathing	Subjective evaluation of breathing difficulty by patient or caregiver, or any of the following: respiratory rate ≥25/min (adults) or SpO2 <90% (unless chronic) or PaO2 <8 kPa or respiratory acidosis	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
sorethroat	Optional	Sore throat		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
deterioration	Optional	Deterioration of general condition (asthenia, loss of weight, anorexia, confusion or dizziness)		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
ili	Obligatory (for outpatient)	Influenza like illness	Fulfilling the EU-ILI case definition	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
sari	Obligatory (for inpatient)	Severe acute respiratory infection	Fulfilling the I-MOVE+ SARI case definition	Numeric (Categorical	0=No 1=Yes 9999=No information	1
respinfection	Optional	Does the patient have a suspected respiratory infection?		Numeric (Categorical)	0=No 1=Yes 9999=No information	1

hosp48h	Obligatory	Was the patient previously hospitalised < 48 hours prior to ILI onset leading to the current hospitalisation?		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
contra	Obligatory	Any contraindication for influenza vaccination	Based on locally used criteria.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
consent	Obligatory	Consent given		Numeric (Categorical)	0=No 1=Yes 999=Not applicable	1
consentkin	Obligatory	Consent given by family member (or alternatively tutor, where applicable)		Numeric (Categorical)	0=No 1=Yes 999=Not applicable	1
comm	Optional	Whether communication with the patient OR consent from next of kin was possible.		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
inst	Obligatory	Institutionalized at the onset of the symptoms	Living in a residence or nursing home (any such institution where nurse present 24/7)	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

prevflu	Obligatory	Did the patient have a previous lab-confirmed influenza in this season?	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
labvirus1	Obligatory	Laboratory result: virus type	Numeric (Categorical)	0=None 1=A 2=B 3=Other influenza not specified 4=Other virus 9999=No information	2
labsubtype1	Obligatory	Laboratory results: virus subtype	Numeric (Categorical)	0=None 1=A(H1N1)pdm09 2=A(H3N2) 3=B Yamagata 4=B Victoria 5=Other influenza 9=Other virus 9999=No information	3
labvirus2	Optional	Laboratory results: virus type (co-infection)	Numeric (Categorical)	0=None 1=A 2=B 3=Other influenza not specified 4=Other virus 9999=No information	1

labsubtype2	Optional	Laboratory results: virus subtype (co-infection)	Numeric (Categorical)	0=None 1=A(H1N1)pdm09 2=A(H3N2) 3=B Yamagata 4=B Victoria 5=Other influenza 9=Other virus 9999=No information	1
seasvaccany	Obligatory	Received influenza vaccination in current season	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
seasvaccbrand	Obligatory	Vaccine brand	Text		Vaxigrip tetra
seasvaccdate	Obligatory	Date of influenza vaccination in 2017-2018	dd/mm/yyyy		11/01/2018
vaccsource	Obligatory	Source of vaccination information	Numeric (Categorical)	1=Registered information (medical record, vaccination card, registered in the system) 2=Recall (from the patient, relative or the health-care professional) 9999=No information	1

seasvaccn1	Optional	Received influenza vaccination in previous season (season n – 1)	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
seasvaccn2	Optional	Received influenza vaccination in season n – 2	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
seasvacckid1	Obligatory	Did the kid (< 9 years) receive 1st dose of influenza vaccination in current season?	Numeric (Categorical)	0=No 1=Yes 999=Not applicable 9999=No information	999
seasvacckid2	Obligatory	Did the kid (<9 years) receive 2nd dose of influenza vaccination in current season?	Numeric (Categorical)	0=No 1=Yes 999=Not applicable 9999=No information	999
seasvaccbrand1	Only if Seasvacckid1 is 1	Vaccine brand	Text		Vaxigrip tetra
seasvaccbrand2	Only if Seasvacckid2 is 1	Vaccine brand	Text		Vaxigrip tetra
seasvaccdate1	Only if Seasvacckid1 is 1	Date of 1st dose of influenza vaccination in the current season (only if Seasvacckid1=1)	dd/mm/yyyy	≥Date within the study period	11/01/2018

seasvaccdate2	Only if Seasvacckid2 is 1	Date of 2nd dose of influenza vaccination in the current season (only if Seasvacckid2=2)		dd/mm/yyyy	≥Date within the study period	11/01/2018
pneumovac	Optional	Received any pneumococcal vaccination	Any time.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
pneumovaccdat	Optional	Date of pneumococcal vaccination	Latest dose.	dd/mm/yyyy		11/01/2018
chronic	Obligatory	Does the patient have at least one chronic disease?	Including obesity (BMI ≥30). Not including smoking or pregnancy.	Numeric (Binary)	0=No 1=Yes 9999=No information	1
liverdis	Optional	Chronic liver disease	Any of the following dg codes (ICD-10): B18, K70-74, K75.0-75.1, K75.3-75.9, K76-77 INCLUDING: Alcoholic liver disease, Toxic liver disease, Hepatic failure, Chronic hepatitis (viral & other), Fibrosis and cirrhosis of liver, Other inflammatory liver diseases, Other diseases of liver EXCLUDING: Clinically insignificant liver cysts	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
diabetes	Optional	Diabetes	Any of the following dg codes (ICD-10): E10-E14, O24 INCLUDING: Any form of diabetes, including sequelae & DM in pregnancy	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

cardiovasc	Optional	Cardiovascular diseases	Any of the following dg codes (ICD-10): A52.0, B37.6, I01-02, I05-09, I11.0, I13.0, I13.2, I20-25, I26-28, I30-43, I44-46, I48, I49.0, I49.5, I50-52, I70-71, Q20-Q28 INCLUDING: all conditions of heart & large vessels that are chronic or likely to have chronic sequelae. Cardiovascular syphilis, endo-, myo- and pericarditis, rheumatic fever, chronic rheumatic heart diseases, congenital malformations, hypertensive (renal) diseases with heart failure, ischaemic heart diseases, diseases of pulmonary circulation, atherosclerosis, cardiomyopathies, most conduction disorders, heart failure, aortic aneurysms & dissecation, other heart diseases and their complications. EXCLUDING: uncomplicated hypertension, previous uncomplicated pulmonary embolism (with no lasting cardiac insufficiency), paroxysmal tachycardias, most cases of premature depolarization.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1

cancer	Optional	Cancer	Any of the following dg codes (ICD-10): C00-97, D37-48, Z85, Z92.3, Z92.6. INCLUDING: All malignant neoplasms (both solid and haematologic) with potential to metastasize, either in treatment, active followup, or <5 years post curative treatment. EXCLUDING: Benign & in situ neoplasms. Basal cell carcinomas. Any cancer previously treated with curative intent & in complete remission for ≥5 years.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

immuno	Optional	Immunodeficiency or organ transplant	Any of the following dg codes (ICD-10): B20-B24, D80–84, D89, Z94 INCLUDING: HIV infections, immunodeficiencies & organ transplants. or iatrogenic: ≥2 week systemic treatment, in the 3 months preceding symptom onset, with any of the following: corticosteroid (≥20 mg prednisolone daily or equivalent), ciclosporin, tacrolimus, mycophenolate, methotrexate, azathioprine, TNF-α blockers and other biological or cytostatic drugs with immunosuppressive effect EXCLUDING: Disorders of the immune system which do not lead to immunosuppression (e.g. some autoimmune conditions).	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

lungdis	Optional	Lung disease	Any of the following dg codes (ICD-10): A15-16, A19, A31.0, B33.4, E84.0, J40-47, J60-70, J80-84, J85-86, J90-91, J92.9, J93-94, J95-99 INCLUDING: TB (pulmonary, miliary but not that of other systems), atypical mycobacteria, cystic fibrosis, asthma, COPD, bronchiectasis and other chronic sequelae of infections, chronic lung diseases due to external agents, interstitial lung diseases, pleural diseases, respiratory failure. EXCLUDING: acute respiratory infections, lung cancer, diseases of pulmonary circulation, pleural plaques without asbestos, previous uncomplicated pneumothorax.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
anemia	Optional	Anemia	Any of the following dg codes (ICD-10): D50-D64 diagnosed before the onset of symptoms. EXCLUDING: coagulopathies, uncomplicated hypersplenism, hepato/splenomegaly (D65-69, D70-77, D80-84, D86, D89)	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
rendisease	Optional	Renal disease	Any of the following dg codes: (ICD-10): I12-13, M10.30, N00-19, N20.0, N25-27, N28.0, N28.9, Q63.9, Z90.5 EXCLUDING: Clinically nonsignificant kidney cysts	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

dement	Optional	Dementia	Any of the following dg codes (ICD-10): F00-03, F05.1, G30-31 EXCLUDING delirium w/o underlying dementia, hydrocephalus.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
stroke	Optional	History of stroke	Any of the following dg codes (ICD-10): I61-64, I67.8, I69, G93.1 INCLUDING: both ischaemic and haemorrhaegic strokes and anoxic brain damage. Also counting previous episodes and clear ischaemic findings seen in cranial imaging (even if fully recovered / no symptoms).	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
rheumat	Optional	Rheumatologic diseases	Any of the following dg codes: ICD-10: M05–09, M13, M30–36, M45 INCLUDING rheumatoid diseases with presumed autoimmune origin and primarily musculoskeletal presentation. EXCLUDING: arthrosis, gout, scoliosis, infectious conditions etc.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
obesity	Optional	Obesity	BMI ≥30 or the dg codes (ICD-10): E66, E68 EXCLUDING: local adiposity and "other hyperalimentation" (=vitamin overdoses etc.)	Numeric (Categorical)	0=No 1=Yes 9999=No information	0

childrisk	Optional	In children: Any perinatal or congenital risk factor?		Numeric (Categorical)	0=No 1=Yes 9999=No information	9999
nhosp	Obligatory	Number of hospitalizations in the last 12 months	Any overnight stay in hospital. (One disease episode counts as one hospitalization even if a patient is moved from one unit to another)	Numeric	0=0 1=1 2=2 3=3 4=4 5=5 6=6 or more 9999=No information	2
gpvisit	Obligatory (for GP studies)	Number of GP consultations in the last year	Any consultation to nurse/GP/specialist in a primary care setting. Not counting follow-up visits for the same cause.	Numeric	≥0 or 9999=No information	5
antiviral	Optional	Has the patient received an antiviral treatment within the 2 weeks before swabbing?		Numeric (Categorical)	0=No 1=Yes 9999=No information	1
statin	Optional	Statin use	At the time of vaccination.	Numeric (Categorical)	0=No 1=Yes 9999=No information	1
pregnancy	Obligatory	Pregnancy	Any trimester at symptom onset.	Numeric (Categorical)	0=No 1=Yes 9999=No information	0
hcw	Optional	Is the patient a healthcare worker?		Numeric (Categorical)	0=No 1=Yes 9999=No information	0

siblings	Optional	(In children) Number of siblings		Numeric	≥0 or 9999=No information	2
bmi	Optional	Body Mass Index	When decimals are provided, please avoid commas	Numeric	10 to 55 or 9999=No information	22.4
smoking	Optional	Smoking status (cigarettes, cigars, pipe, hookah). Not counting exclusively chew tobacco or snus.	Never-smoker: <100 cigarettes during their lifetime. Ex-smoker: has smoked ≥100 cigarettes over lifetime but has stopped ≥3 months ago. Occasional smoker: has smoked ≥100 cigarettes over lifetime and has still smoked in the 3 months preceding symptom onset, but not daily. Daily smoker: has smoked ≥100 cigarettes over lifetime and smokes daily.	Numeric (Categorical)	0=Never-smoker 1=Ex-smoker 2=Occasional smoker 3=Daily smoker 9999=No information	0
functstatus	Optional	Dependency / Patient has difficulty in at least 1 of these categories: bathing dressing eating going to the toilet stairs walk wheelchair user	Difficulty = needs help from others	Numeric (Categorical)	0=No 1=Yes 999=Not applicable 9999=Unknown	0

2 Annex 2. DRIVE Electronic Study Support Application (ESSA): user manual

2.1 General introduction and description of the ESSA

In the DRIVE project, influenza vaccine effectiveness (IVE) data from several independently operating national and regional study sites will be analysed together to obtain brand-specific IVE estimates.
The DRIVE Electronic Study Support Application (ESSA) facilitates the quality control of the data and the safe uploading of data to the central DRIVE research server. The DRIVE ESSA is a web application accessible from any web browser serving the following purposes:

- Aiding research sites to do the quality assurance of their data by automatically performing data quality checks
- Providing a visual summary of the data
- Allowing research sites to share the visual summaries and tables for monitoring purposes
- Allowing research sites to safely upload their data to the central DRIVE Research Server for analysis

For any questions, please contact: maria.alexandridou@p-95.com

2.2 ESSA Login process

In order to access to the DRIVE ESSA, use the following link: **http://apps.p-95.com/essa1/**.
The user can log in to the web application by using a personal email address and password combination.

2.3 Data Flow

In Figure 1 the data flow to the DRIVE Research server is detailed. The interim and final study data is uploaded by the DRIVE research study sites to the ESSA Server using a secure connection. After the data is checked, the DRIVE research study site can decide to share minimal data for monitoring or to transfer the final data to the DRIVE Research Server for analysis. All data transfers will take place using a secure connection.

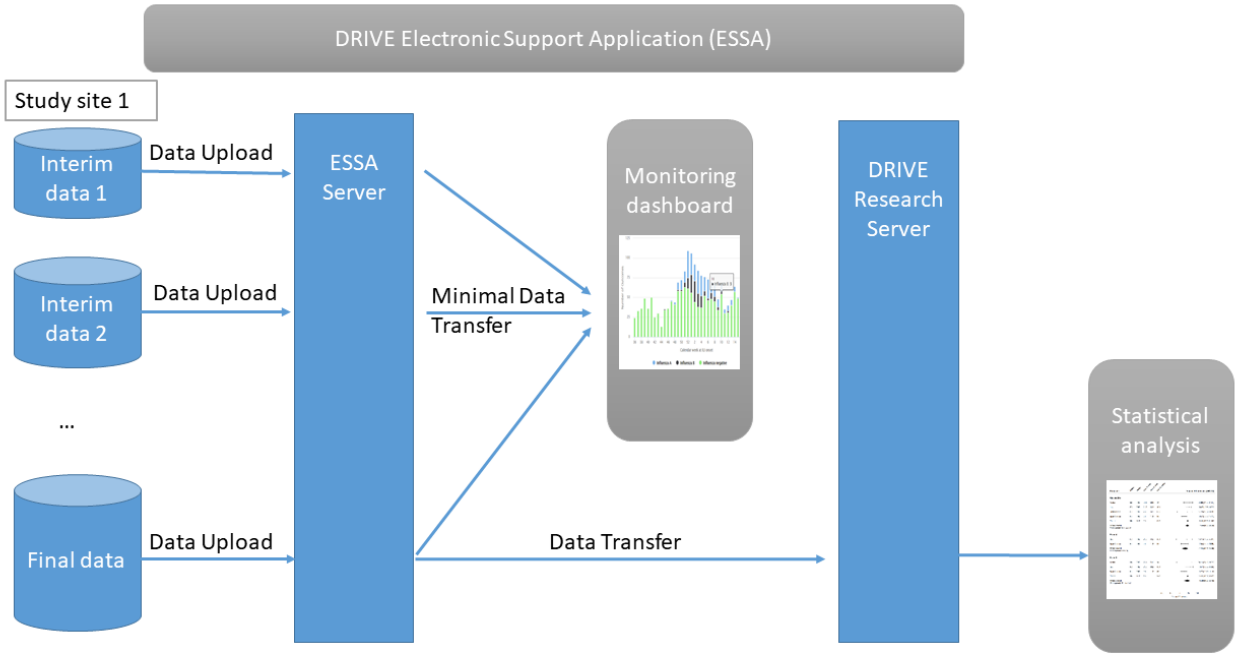


Figure 1 Data flow

2.4 Main commands of the application

The sidebar on the left of the ESSA is the main control tool of the application. The options available in the sidebar are:

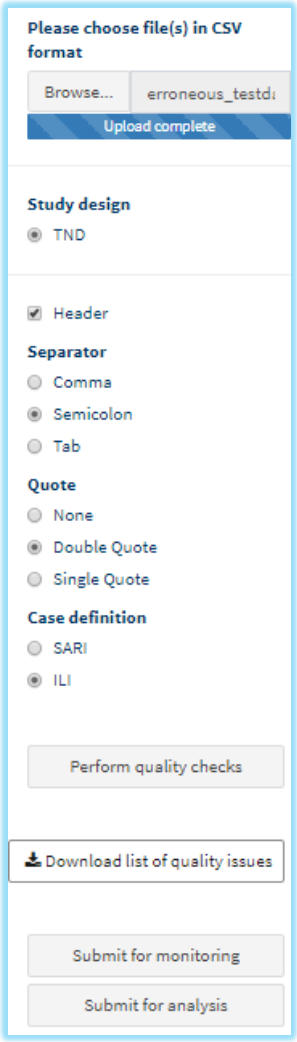


Figure 2 ESSA control side bar

- The button **Browse** to select the dataset for uploading to the ESSA;
- The **Study design** button to select the study design of the uploaded dataset. Currently, only TND (Test-Negative Design) option is available;
- The **Header** checkbox to indicate if the first row of the data contains the column headers;
- The radio buttons **Separator** to indicate whether the data fields are separated by a comma, semicolon or tab;
- The radio buttons **Quote** to indicate whether each field in the data is contained within single quotes, double quotes or no quotes are used;
- The radio buttons **Case definition** to indicate whether the SARI (Severe Acute Respiratory Infection) or ILI (Influenza Like Illness) case definition was used for the uploaded dataset;
- The action button **Perform quality checks** to perform the quality checks of the data;
- The action button **Download list of quality issues** to download a list with quality issues found through performing the quality checks.
- The action button **Submit for monitoring** to allow sharing the data summaries (tables and figures) with the DRIVE study team;
- The action button **Submit for analysis** to indicate the uploaded data can be used for analysis;

2.5 Step 1: upload a dataset to the ESSA

2.5.1 Dataset format

Datasets must be in CSV format (with no size limit) and be compliant with the minimum data requirements (link). Some variables are obligatory, other optional. The compliance to the minimum data requirements will be checked by the ESSA application.

2.5.2 Upload data procedure

1. Click the **Browse** button located at the top of the Control Sidebar of the ESSA (Figure 2**Fout! Verwijzingsbron niet gevonden.**).

2. Browse computer folders and select the .csv file to upload. The file will be uploaded to the ESSA.

3. Indicate how the data should be read using the **Study design**, **Header**, **Separator**, **Quote** and **Case definition** buttons.

4. Preview the uploaded data and change the options above if the data are not properly read. See examples below.

An example of a properly uploaded dataset is given in Figure 3. An example of data that is not properly uploaded is given in Figure 4. In this example view, the following problems are present and can be solved using the options available in the ESSA control sidebar:

Problem	How to solve
Variables names are in the first data row of the dataset	Check the **Header** checkbox in the sidebar
Variables are not placed in columns	Select among the **Separator** options, the proper field separator (“Comma” in this case)
Variables values appear in quotes	Select among the **Quote** options, how the fields are contained (“Double Quote” in this case)

Show 10 entries

X	idcountry	id	region	gp	hosp	sex	dob	onsetdate
1	SP	1000001	Valencia		10	1	15/11/1949	07/09/2017
2	SP	1000002	Valencia		10	0	01/04/1923	08/09/2017
5	SP	200003	Valencia		2	1	18/09/1942	09/09/2017
6	SP	200005	Valencia		2	0	30/06/1971	05/09/2017
12	SP	700005	Valencia		7	1	23/01/1975	09/09/2017
14	SP	1000007	Valencia		10	1	08/12/1940	07/09/2017
16	SP	1000009	Valencia		10	0	11/04/1934	10/09/2017
22	SP	200010	Valencia		2	0	04/02/1934	07/09/2017

Figure 3 View of data that is properly uploaded

Show 10 entries

itkid","comm","inst","prevflu","resident","labvirus","labsubtype","seasvaccany","seasvaccbrand","sea

"1","SP",1000001,"Valencia",NA,10,1,"15/11/1949","07/09/2017","11/09/2017","10/09/2017",1,"25/09,

"2","SP",1000002,"Valencia",NA,10,0,"01/04/1923","08/09/2017","11/09/2017","10/09/2017",0,"

"5","SP",200003,"Valencia",NA,2,1,"18/09/1942","09/09/2017","11/09/2017","10/09/2017",0,"",

"6","SP",200005,"Valencia",NA,2,0,"30/06/1971","05/09/2017","11/09/2017","10/09/2017",0,"",0,

"12","SP",700005,"Valencia",NA,7,1,"23/01/1975","09/09/2017","11/09/2017","11/09/2017",0,"",1

"14","SP",1000007,"Valencia",NA,10,1,"08/12/1940","07/09/2017","12/09/2017","11/09/2017",0,"

"16","SP",1000009,"Valencia",NA,10,0,"11/04/1934","10/09/2017","12/09/2017","11/09/2017",0,"

Figure 4 View of data that is not properly uploaded

2.6 Step 2: Perform the data quality checks

Data quality checks are performed upon pressing the action button **Perform quality checks** (Figure 2).
The results of the quality checks will be summarized on the screen. See Figure 5 and Figure 6 for the data quality summaries of data without and with data quality issues.

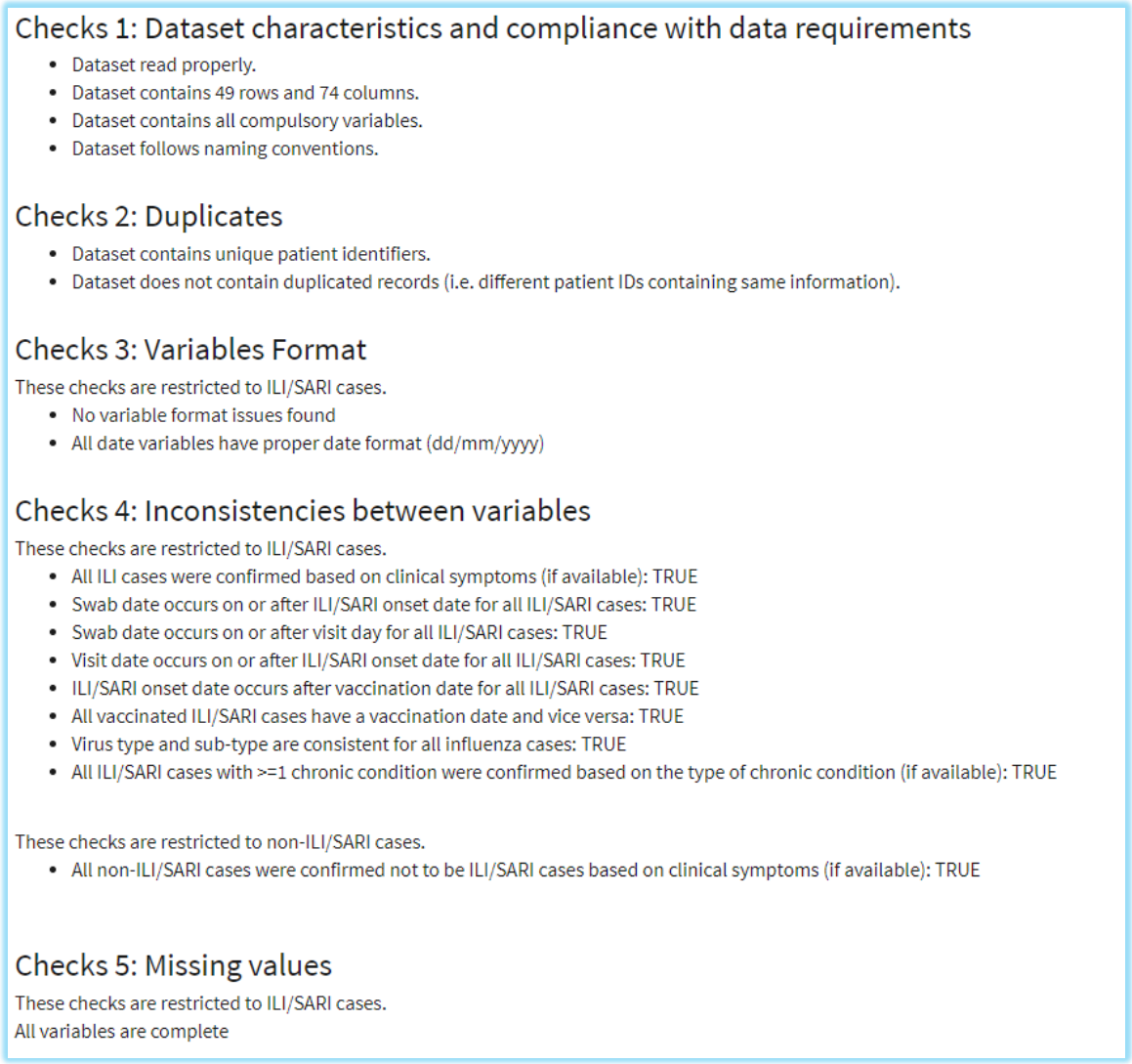
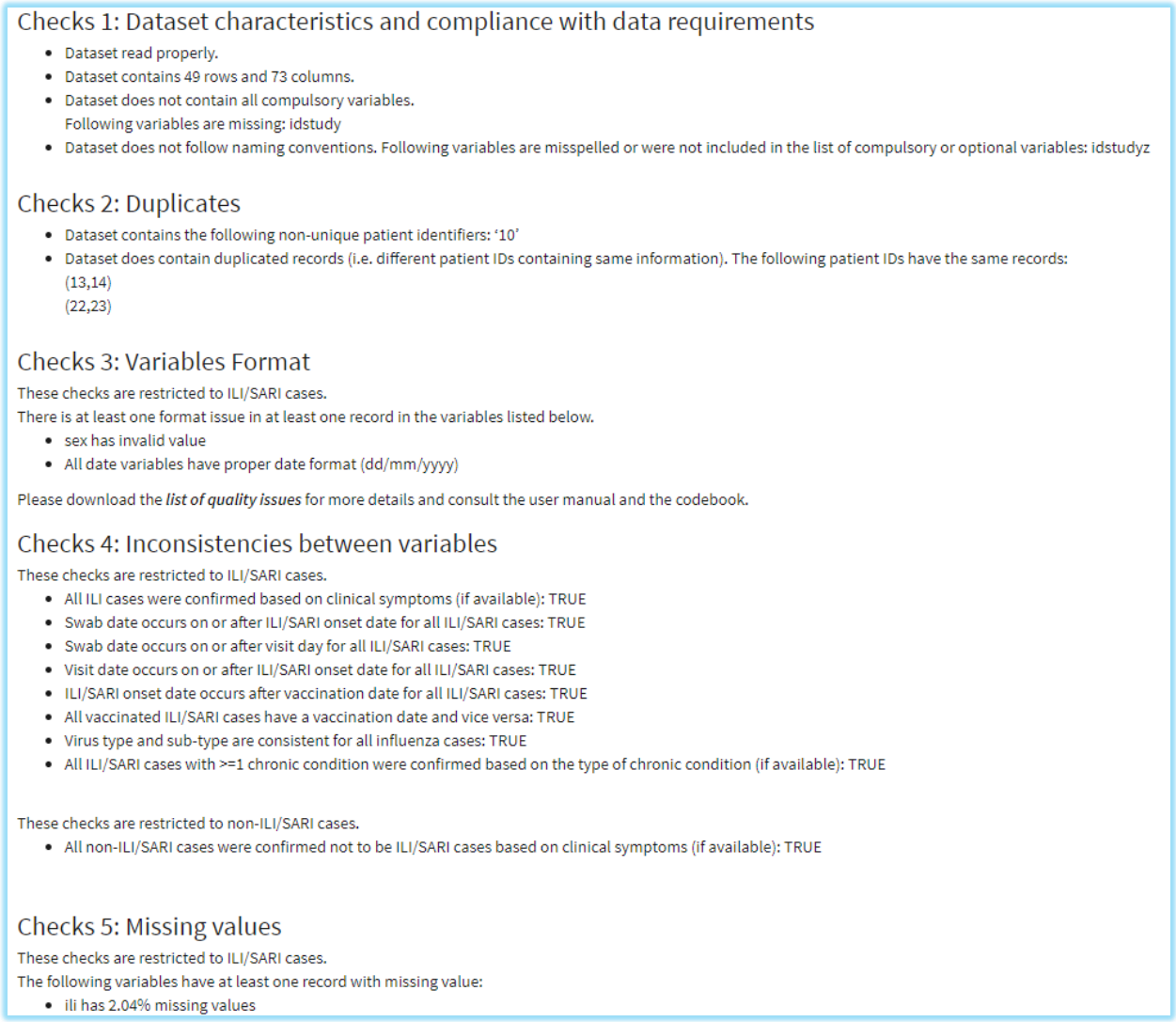


Figure 5 Data quality summary: data without quality issues



Checks 1: Dataset characteristics and compliance with data requirements

- Dataset read properly.
- Dataset contains 49 rows and 73 columns.
- Dataset does not contain all compulsory variables.
  Following variables are missing: idstudy
- Dataset does not follow naming conventions. Following variables are misspelled or were not included in the list of compulsory or optional variables: idstudyz

Checks 2: Duplicates

- Dataset contains the following non-unique patient identifiers: '10'
- Dataset does contain duplicated records (i.e. different patient IDs containing same information). The following patient IDs have the same records:
  (13,14)
  (22,23)

Checks 3: Variables Format

These checks are restricted to ILI/SARI cases.
There is at least one format issue in at least one record in the variables listed below.

- sex has invalid value
- All date variables have proper date format (dd/mm/yyyy)

Please download the ***list of quality issues*** for more details and consult the user manual and the codebook.

Checks 4: Inconsistencies between variables

These checks are restricted to ILI/SARI cases.

- All ILI cases were confirmed based on clinical symptoms (if available): TRUE
- Swab date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE
- Swab date occurs on or after visit day for all ILI/SARI cases: TRUE
- Visit date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE
- ILI/SARI onset date occurs after vaccination date for all ILI/SARI cases: TRUE
- All vaccinated ILI/SARI cases have a vaccination date and vice versa: TRUE
- Virus type and sub-type are consistent for all influenza cases: TRUE
- All ILI/SARI cases with >=1 chronic condition were confirmed based on the type of chronic condition (if available): TRUE

These checks are restricted to non-ILI/SARI cases.

- All non-ILI/SARI cases were confirmed not to be ILI/SARI cases based on clinical symptoms (if available): TRUE

Checks 5: Missing values

These checks are restricted to ILI/SARI cases.
The following variables have at least one record with missing value:

- ili has 2.04% missing values

Figure 6 Data quality summary: data with quality issues

2.6.1 Checks 1: Dataset characteristics and compliance with data requirements

The ESSA will provide an indication about the status of the upload and inform the user about the number of rows and columns processed in the uploaded data file.

The data will be validated with respect to the minimum data requirements. If there are compulsory variables missing in the variable names, the ESSA will provide an error message ("Dataset does not contain all compulsory variables") and provide a list of missing variables in the data uploaded. The ESSA will also check if the variable names are consistent with respect to the naming conventions specified in the DRIVE codebook. It will indicate any misspelt or redundant variable names.

2.6.2 Checks 2: Duplicates

The ESSA will check whether there are duplicated patient IDs (patient identifiers) and whether different patient IDs contain the same information (i.e. all the information is the same apart from the patient IDs). If there are any duplicated patient IDs, the DRIVE ESSA will show the duplicated patient IDs. If there are multiple patient IDs that contain the same information, the DRIVE ESSA will show these multiple patient IDs sharing the same information between round brackets.

2.6.3 Checks 3: Variables Format

The variables are checked with respect to the acceptable format and value ranges as specified in the DRIVE codebook. For example, valid values for the variable “fever” are 0 (no), 1 (yes) and 9999 (no information). Other values will generate a data quality issue and will be mentioned in the data quality summary. Details on the variable format issues will be given in the CSV file that can be downloaded using the action button in the sidebar **Download list of quality issues.**

2.6.4 Checks 4: Inconsistencies between variables

Several consistency checks between variables are performed. The different consistency checks are described in Table 2 and Table 3. If a check cannot be done because variables are not included in the data set, then the message ‘Check not possible’ will be given. Details on the inconsistencies found will be given in the CSV file that can be downloaded using the action button in the sidebar **Download list of quality issues.**

Table 2 Inconsistency checks for the ILI/SARI cases only

All ILI cases were confirmed based on clinical symptoms (if available): TRUE/FALSE

(activates only if the ILI case definition is chosen on the side bar)

Checks if every ILI patient does follow the case definition (Sudden onset & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are ILI patients whose symptoms do not indicate that they have ILI, then this is displayed on the data quality summary as FALSE. If the ILI symptoms are not included in the submitted data set, then the default message is TRUE.

All SARI cases were confirmed based on clinical symptoms (if available): TRUE/FALSE

(activates only if the SARI case definition is chosen on the side bar)

Checks if every SARI patient does follow the case definition (Deterioration & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are SARI patients whose symptoms do not indicate that they have SARI, then this is displayed on the data quality summary as FALSE. If the SARI symptoms are not included in the submitted data set, then the default message is TRUE.

Swab date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of swab is on the same date or after the date of onset.

Swab date occurs on or after visit day for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of swab is on the same date or after the date of visit.

Visit date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of visit is on the same date or after the date of ILI/SARI onset.

ILI/SARI onset date occurs after vaccination date for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of ILI/SARI onset is on the same date or after the date of vaccination.

All vaccinated ILI/SARI cases have a vaccination date and vice versa: TRUE/FALSE

Checks whether the vaccinated persons have a vaccination date and vice versa.

Virus type and sub-type are consistent for all influenza cases: TRUE/FALSE

Checks if virus type and virus subtype are consistent. (e.g. a patient with influenza A and subtype B Yamagata would be flagged as inconsistent)

All ILI/SARI cases with >=1 chronic condition were confirmed based on the type of chronic condition (if available) : TRUE/FALSE

Checks if the chronic case definition can be confirmed using at least one of the following conditions: chronic liver disease, diabetes, cardiovascular diseases, cancer, immune-deficiencies, lung disease, anaemia, renal disease, dementia, stroke, rheumatoid disease, obesity, perinatal/congenital risk factor. If chronic symptoms are not included in the submitted data set, then the default message is TRUE.

Table 3: Inconsistency checks for patients that are non-ILI/SARI cases

All non-ILI cases were confirmed not to be ILI cases based on clinical symptoms (if available): TRUE/FALSE Checks if every non-ILI patient does not follow the ILI case definition (Sudden onset) & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are non-ILI patients whose symptoms indicate that they have ILI, then this is displayed on the screen as false. If the ILI symptoms are not included in the submitted data set, then the default message is true.
All non-SARI cases were confirmed not to be SARI cases based on clinical symptoms (if available): TRUE/FALSE Checks if every non-SARI patient does not follow the SARI case definition (Deterioration) & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are non-SARI patients whose symptoms indicate that they have SARI, then this is displayed on the screen as false. If the SARI symptoms are not included in the submitted data set, then the default message is true.

2.6.5 Checks 5: Missing values

The ESSA checks for each variable the amount of missing data. For each variable with missing data, the percentage of missing data will be given in the data quality summary.

2.7 Step 3: Download the list with quality issues found

Upon performing the quality checks, a list with identified quality issues will be automatically generated. A CSV file with the list of identified quality issues can be downloaded by pressing the action button in the sidebar **Download list of quality issues**. Figure 7 gives an excerpt of such a CSV file with identified quality issues. A record is generated for every patient ID for which at least one quality issue was found.

ID	issues
1	Row issue(s): - swabdate occurs earlier than onsetdate
2	Row issue(s): - agemonths has invalid value
3	Row issue(s): - agemonths has invalid value
7	Row issue(s): - ILI case definition cannot be verified based on symptoms
8	Row issue(s): - bmi has negative value
12	Row issue(s): - virus type not consistent with subtype
14	Row issue(s): - idcountry is not unique
17	Row issue(s): - onsetdate occurs earlier than vaccination date - virus type not consistent with subtype
20	Row issue(s): - death has invalid value
21	Row issue(s): - smoking has invalid value
23	Row issue(s): - vaccination status not consistent with vaccination date
28	Row issue(s): - chronic case definition inconsistent with symptoms
32	Row issue(s): - vaccination status not consistent with vaccination date
33	Row issue(s): - consentkin has invalid value
49	Row issue(s): - setting has invalid value - seasvaccn2 has invalid value

Figure 7 CSV file with quality issues

2.8 Step 4: Explore data summaries

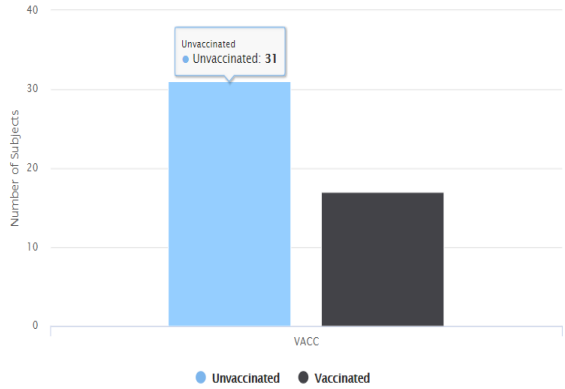
If the required information is provided, the ESSA generates three tables; one giving information on the number of subjects; one giving information on the number of brands and one giving information on the number of influenza cases by age group (Figure 8). Graphical summaries of the data are provided as well, regarding exposure (Figure 9), the outcome variables (Figure 10) and covariates (Figure 11).

	N
Nr. Subjects	49
Nr. Influenza cases	32
Nr. Vaccinated subjects	17

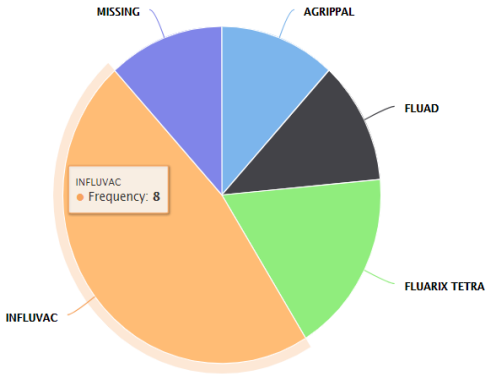
Brand	Nr. Subjects
AGRIPPAL	2
FLUAD	4
FLUARIX TETRA	3
INFLUVAC	8

Age	Nr. Influenza Cases
Children (0-14y)	10
Adults (15-64y)	14
Elderly (65+y)	8

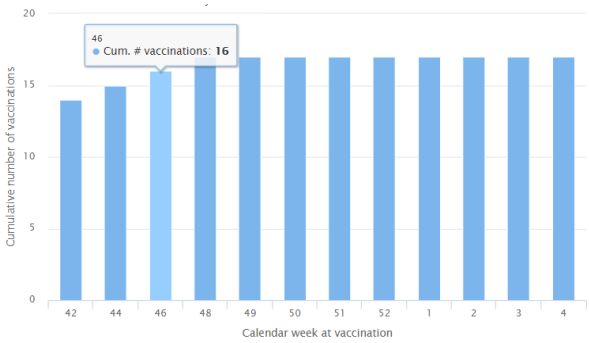
Figure 8 Data summary tables provided by ESSA



Number of (un)vaccinated subjects

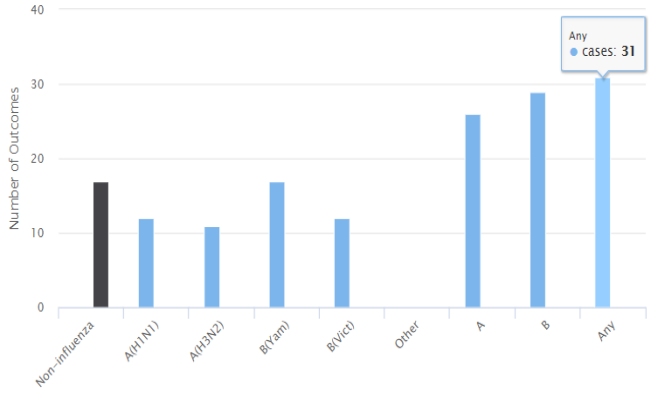


Distribution of vaccine brands

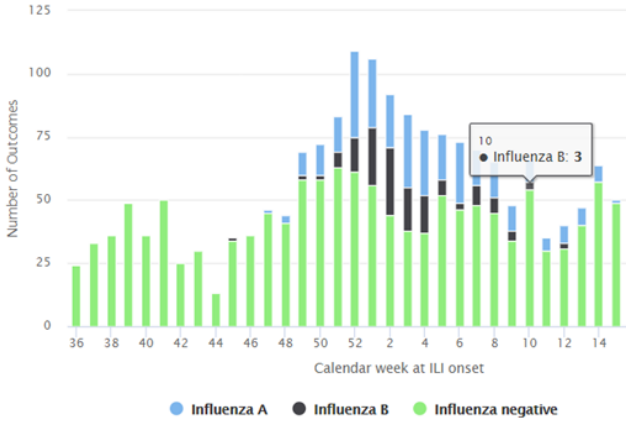


Cumulative number of vaccinated subjects over time

Figure 9 Visualizations regarding exposure

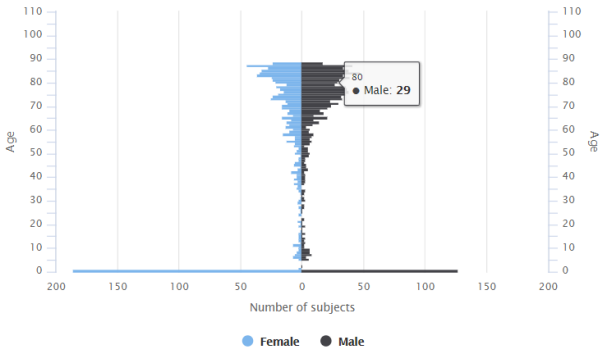


Number of cases and controls with laboratory confirmed influenza by type and subtype

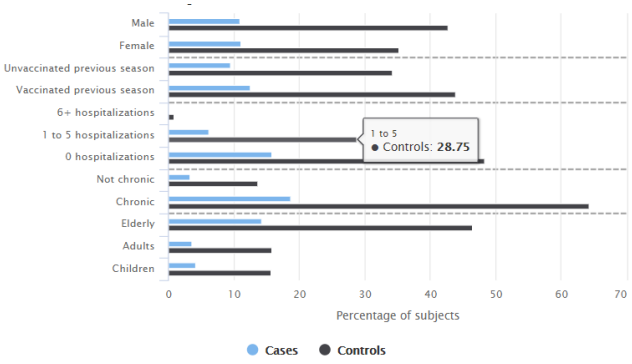


Number of laboratory-confirmed influenza infections by type over time

Figure 10 Visualizations regarding the outcome variables



Age and gender pyramid



Distribution of covariates among cases and controls

Figure 11 Visualization regarding covariates

2.9 Step 5: Share data for monitoring

After performing the quality checks and possibly modifying the data, the data summary (Tables and Figures as described in Step 4) can be shared with the DRIVE study team for data monitoring purposes. For sharing the tables and figures, press the action button **Submit for monitoring.** Sharing of data summaries for monitoring purposes can be repeatedly done throughout the influenza season.

The summary tables and figures shared by all study sites are accessible from **http://apps.p-95.com/essa2/**. The information is only accessible to members of the DRIVE study team.

2.10 Step 6: Submit data for analysis

To submit the final and clean data for analysis, press the action button **Submit for analysis.**

3 Annex 3. List of chronic conditions

Table S3.1: Data collected on chronic conditions – test-negative design primary care studies, 2018-2019

Country	Austria	Italy	Italy	Spain	UK
Site	MUV	CIRI-IT	ISS	RS	RCGP/UNIS
Chronic liver disease	Yes	Yes	No	Yes	Yes
Diabetes	Yes	Yes	No	Yes	Yes
Cardiovascular diseases	Yes	Yes	No	Yes	Yes
Cancer	Yes	Yes	No	Yes	Yes
Immunodeficiency or org an transplant	Yes	Yes	No	Yes	Yes
Lung disease	Yes	Yes	No	Yes	Yes
Anemia	Yes	Yes	No	Yes	Yes
Renal disease	Yes	Yes	No	Yes	Yes
Dementia	Yes	Combined with stroke	No	Yes	Yes
History of stroke	Yes	Combined with dementia	No	No	Yes
Rheumatologic diseases	Yes	Yes	No	No	Yes
Obesity (BMI≥30)	No	Yes	No	BMI≥40 in adults, ≥35 in adolescents, ≥30 in children	Yes
In children: Any perinatal or congenital risk factor?	No	Yes	No	No	Yes
Other	n/a	(1)	n/a	n/a	n/a

BMI: body mass index

(1) Metabolic disease, hematopoietic organs and hemoglobinopathies, Chronic inflammatory disease and intestinal malabsorption, Nutritional deficiency, Leukemia, lymphoma, Pathologies associated with an increased risk of aspiration of respiratory secretions

Table S3.2: Data collected on chronic conditions – test-negative design hospital studies, 2018-2019

Country	Austria	Finland	Italy	Romania	Spain	Spain
Site	MUV	HUCH	BIVE-HOSP	NIID	FISABIO	VHUH
Chronic liver disease	Yes	Yes	Yes	No	Yes	Yes
Diabetes	Yes	Yes	Yes	No	Yes	Yes
Cardiovascular diseases	Yes	Yes	Yes	No	Yes	Yes
Cancer	Yes	Yes	Yes	No	Yes	Yes
Immunodeficiency or organ transplant	Yes	Yes	Yes	No	Yes	Yes
Lung disease	Yes	Yes	Yes	No	Yes	Yes
Anemia	Yes	Yes	Yes	No	Yes	No
Renal disease	Yes	Yes	Yes	No	Yes	Yes
Dementia	Yes	Yes	Yes	No	Yes	No
History of stroke	Yes	Yes	Yes	No	No	No
Rheumatologic diseases	Yes	Yes	Yes	No	Yes	Yes
Obesity (BMI≥30)	No	Yes	Yes	No	Yes	BMI≥40
In children: Any perinatal or congenital risk factor?	No	n/a	No	Yes	No	Yes
Other		-Chronic neurological or neuromuscular diseases	(1)	-Metabolic disorder	(2)	n/a

BMI: body mass index
(1) Metabolic disease, hematopoietic organs and hemoglobinopathies, Chronic inflammatory disease and intestinal malabsorption Nutritional deficiency, Leukemia, lymphoma, Pathologies associated with an increased risk of aspiration of respiratory secretions
(2) Cerebrovascular disease, peripheral arteriopathy, disease of the endocrine system other than diabetes, neuromuscular or neurodegenerative disease.

Table S3.3: Data collected on chronic conditions – register-based cohort study, 2018-2019

Country	**Finland**
Site	**THL**
Chronic liver disease	Yes
Diabetes	Yes
Heart of cardiovascular disease	Yes
Cancer	Yes
Immuno-deficiency or organ transplant	Yes
Lung disease	Yes
Anemia	Yes
Renal disease	Yes
Dementia	Yes
History of stroke	Yes
Rheumatologic disease	Yes
Any perinatal or congenital risk factor in children	No
Obesity	Yes
At least one chronic disease	Yes
At least 5 chronic diseases	No
Other	

Table S3.4: Data collected on chronic conditions – clinical cohort study pregnant women and their infants, 2018-2019

Country	**Greece**
Site	**UoA**
Chronic liver disease	No
Diabetes	Yes
Heart of cardiovascular disease	Yes
Cancer	Yes
Immuno-deficiency or organ transplant	No
Lung disease	Yes (COPD)
Anemia	No
Renal disease	Yes
Dementia	No
History of stroke	No
Rheumatologic disease	No
Any perinatal or congenital risk factor in children	n/a
Obesity	No
At least one chronic disease	Yes
At least 5 chronic diseases	?
Other	

Table S3.5: Data collected on chronic conditions – clinical cohort study healthcare workers, 2018-2019

Country	**Italy**
Site	**CIRI-IT**
Chronic liver disease	Yes
Diabetes	And other metabolic diseases
Heart of cardiovascular disease	Yes
Cancer	No
Immuno-deficiency or organ transplant	Yes
Lung disease	Yes
Anemia	Yes
Renal disease	Yes
Dementia	Combined with stroke
History of stroke	Combined with dementia
Rheumatologic disease	Yes
Any perinatal or congenital risk factor in children	n/a
Obesity	Yes
At least one chronic disease	Yes
At least 5 chronic diseases	?
Other	-hematopoietic organs and hemoglobinopathies -Chronic inflammatory disease and intestinal malabsorption -Pathologies associated with an increased risk of aspiration of respiratory secretions -leukemia, lymphoma -Nutritional deficiency

n/a: not applicable

Annex 3. Data Quality Reports

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: Medical University Vienna, AUSTRIA

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Kaatje Bollaerts (P95) Maria Alexandridou (P95) Nick De Smedt (P95)
Data provider	Monika Redlberger-Fritz (Medical University Vienna, Austria)

Version	Date	Description
V0.1	20 May 2019	First Draft
V0.2	22 May 2019	Changes to first draft
V0.3	24 May 2019	Shared with QCAC + MDR
	27 May 2019	Incorporated comments by Ann-Mary Kirby
V1.0	05 June 2019	Final version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Medical University of Vienna (Austria). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

Medical University Vienna, Austria

The Medical University of Vienna is the national reference laboratory in Austria. They participate to the DRIVE 2018/19 multi-site study with data collected using a test-negative design study. They provided two datasets, one with data collected within the primary care setting and one with data collected in hospitals. A description of the study characteristics can be obtained from the DRIVE SAP 2018/19.

The data sets used for central data quality checks were the hospital and primary care datasets uploaded for 'analysis' using the DRIVE Electronic Study Support Application on 09/05/2019. The uploaded Austria hospital dataset contained records on 35 patients, only one of them received influenza vaccination during the 2018/19 period, for which the brand name was missing. As such, the hospital dataset from Austria was not considered for analysis and no additional data quality checks were done. The uploaded Austria primary care dataset contained records on 1227 patients. The subsequent sections describe the primary care data.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	Yes
#3.	All records included in the dataset are ILI/SARI cases?	Yes
#4.	There are no duplicated records in the dataset?	No, 4 records contained the same information but have different patient identifiers. See Table 2.
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	No, the variable 'seavaccn1' had 4 records with invalid values. See Table 2
#6.	ILI/SARI cases could be verified based on clinical symptoms	No, the information on the symptoms was missing for 99.5% of the records. No data cleaning action can be taken as the information is not collected. This will be taken into account when interpreting the results.
#7.	All records have onset date before swab date?	No, 3 records violated this condition. See Table 2
#8.	For hospitalized cases, all records have hospital admission no longer than 2 days before swab date?	Yes
#9.	All records have visit date after onset date?	No, 3 records violated this condition. See Table 2
#10	All records have vaccination date before onset date?	Yes
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	No, 28 records violated this condition. See Table 2
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	Yes
#14	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	No, 59 records were recorded as 'no chronic condition' whereas 'obesity' was present
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#4	(987,998)	Patient ids appeared to be duplicates. Checked with data site.	These are not duplicates. The data site was contacted and confirmed that both records refer to different persons with different names/dates of birth. Both records are retained in the data
#4	(1236, 1241)	Patient ids appeared to be duplicates. Checked with data site.	Same as above. Both records are retained in the data
#5	133	'seasvaccn1' = 2014 (seasvaccn1 = vaccinated previous season)	The variable 'seasvaccn1' refers to 'being vaccinated in the previous influenza season' (so vaccinated in 2017).

			It seems that the last year at previous vaccination was 2014.The variable was recoded 'seasvaccn1' = 0
#5	256	'seasvaccn1' = 2008	Similar as above. The variable was recoded: 'seasvaccn1' = 0
#5	279	'seasvaccn1' = 2012	'Similar as above. The variable was recoded: 'seasvaccn1' = 0
#5	291	'seasvaccn1' = 2009	Similar as above. The variable was recoded: 'seasvaccn1' = 0
#7, #9	452	Inconsistencies in swab date, onset date and visit date: onset date after swab date/visit date	The site was contacted and the original CRF was verified. There was a typo at data entry of the onset date. This has been corrected.
#7, #9	491	Inconsistencies in swab date, onset date and visit date: onset date after swab date/visit date	Same as above
#7, #9	1171	Inconsistencies in swab date, onset date and visit date: onset date after swab date/visit date	Same as above
#11	75,167,170,229,259,344,368,400,402,531,542,648,678,738,739,887,965,979,1090,1103,1123,1133,1174,1195,1237,1252	Vaccination status is positive, but no vaccination date given	These records will be discarded when applying the study exclusion criteria (see next Section)
#14	21, 37, 46, 61, 64, 70, 71, 77, 82, 95, 115,126, 131, 138, 185, 203, 241, 244 ,252, 266, 270, 276, 280, 286, 321, 344, 365, 373, 377, 381, 413, 415, 424, 432, 441, 461, 473, 478, 481, 499, 504, 524, 555, 570, 598, 599, 635, 707, 735, 772, 824; 829, 891, 1004, 1125, 1131, 1156, 1220	Variable '>1 chronic condition' is inconsistent with variable 'obesity' ('chronic = 0' & 'obese' = 1).	The variable 'chronic' means 'at least 1 chronic condition' out of a list of chronic conditions (including obesity). All these records have been relabelled as 'chronic' = 1

We additionally found two records corresponding to children of 1 year old with a double vaccination. The variable ‘seasvaccbrand’ was used to record this information while the variables ‘seasvaccbrand1’ and ‘seasvaccbrand2’ should have been used. Upon contacting the site, it was confirmed that the patients received two vaccines. Both vaccination dates were also provided by the site. We also identified one record with a visit date after data upload. This was a typo and was corrected. These additional data quality checks and corrections are summarized in Table 3.

Table 3 Additional data quality issues, and corrections

Patient id	Data quality issue	Action
908	Variable ‘seasvaccbrand = ‘Vaxigrip Tetra + Fluarix Tetra’.	The site was contacted, and the following changes were made: ‘seasvaccbrand’= VAXIGRIP TETRA/ FLUARIX TETRA ‘seasvaccdate’ = 14.12.2018 ‘seasvacckid1”= Vaxigrip Tetra ‘seasvacckid2’ = “Fluarix Tetra” ‘seasvaccdate1” = 13.11.2018 Seasvaccdate2” = 14.12.2018
1231	Variable ‘seasvaccbrand = ‘Vaxigrip Tetra + Fluarix Tetra’.	The site was contacted, and the following changes were made: ‘seasvaccbrand’= VAXIGRIP TETRA/ FLUARIX TETRA ‘seasvaccdate’ = 30.11.2018 ‘seasvacckid1”= Vaxigrip Tetra ‘seasvacckid2’ = “Fluarix Tetra” ‘seasvaccdate1” = 24.10.2018 Seasvaccdate2” = 30.11.2018
1202	Visit and swab dates were after data upload	The site was contacted. The original CRF was verified and it appeared to be an error at data entry. The following changes were made: ‘swabdate’ = 13.03.2019 ‘visitdate’ = 13.03.2019

After data cleaning, a total of 1227 records were retained for further processing.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. has symptom onset outside the study period
2. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
3. is less than 6 months of age at the time of the onset of the symptoms
4. has a contraindication for influenza vaccine
5. is institutionalised at the time of symptoms onset
6. will have the respiratory specimen taken ≥ 8 days after ILI onset
7. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

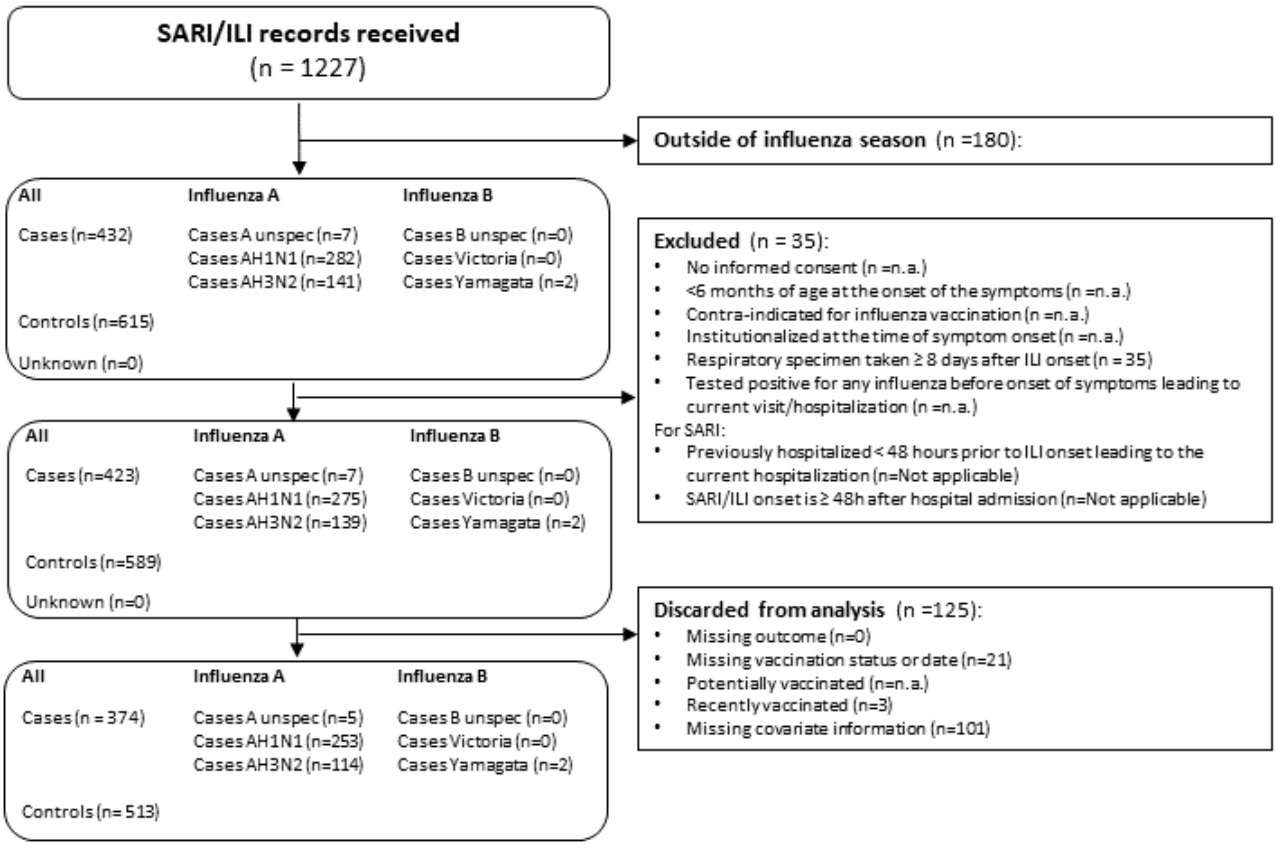


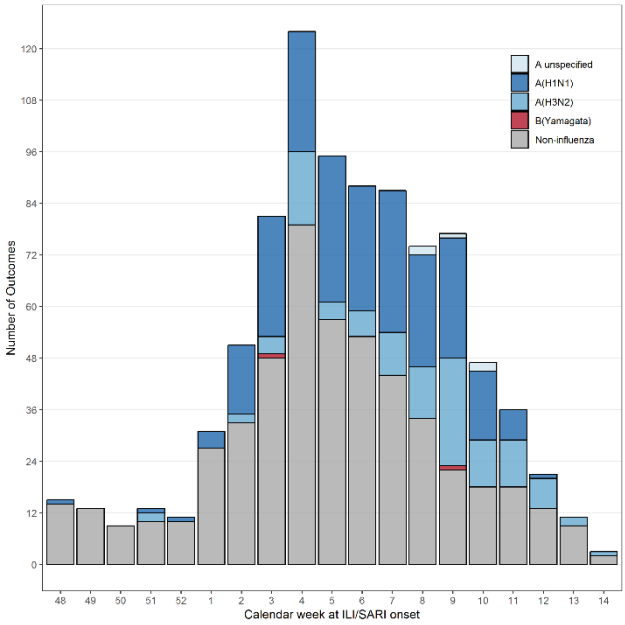
Figure 1 Attrition diagram

Data summary

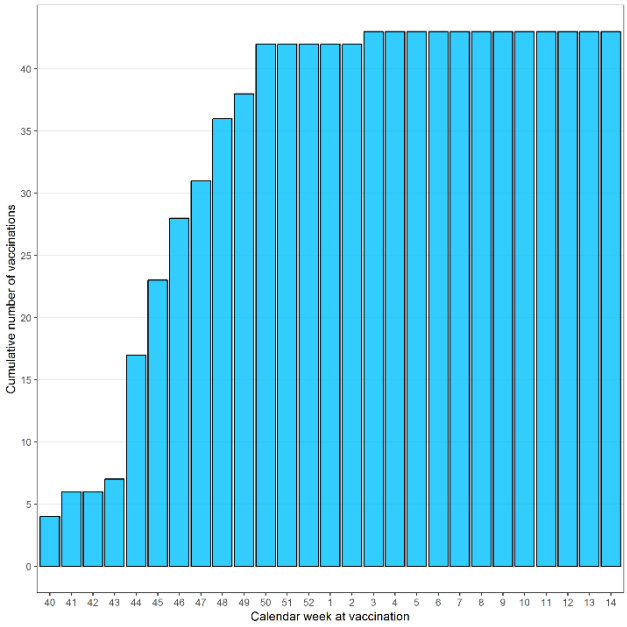
A total of 887 subjects were retained for analysis, with the majority of subjects being either children (<17 yrs) or adults (18-64yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

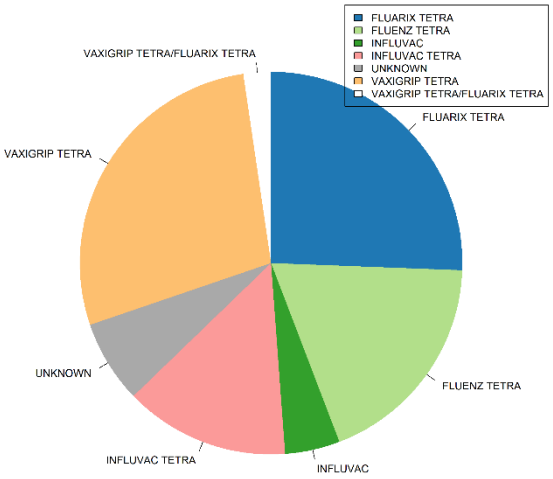
Characteristic	Number of subjects
Subjects	887
Influenza cases	374
Vaccinated subjects	43
Age	
Children (0-17y)	432
Adults (18-64y)	422
Elderly (65+y)	33
Influenza vaccine brand	
Fluarix Tetra	11
Fluenz Tetra	8
Influvac	2
Influvac Tetra	6
Vaxigrip Tetra	12
Vaxigrip Tetra/Fluarix Tetra	1
Unknown	3



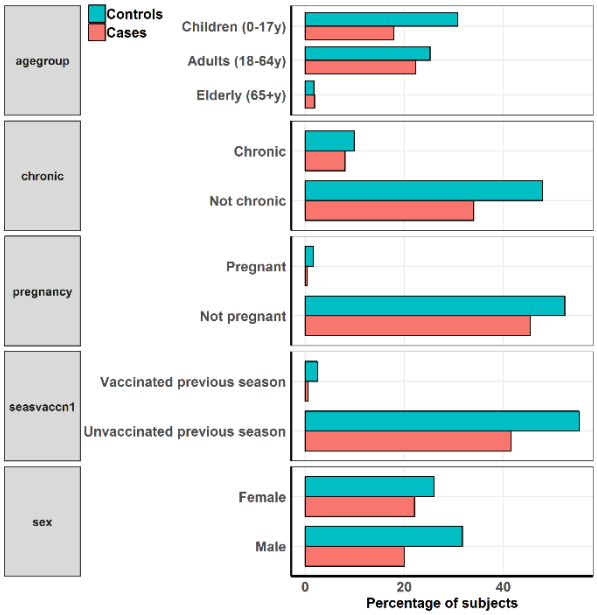
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili, ITALY

TEST NEGATIVE DESIGN STUDY

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Maria Alexandridou (P95) Kaatje Bollaerts (P95) Nick De Smedt (P95)
Data provider	Stefano Mosca

Version	Date	Description
V0.1	27 May 2019	First Draft
V1.0	05 June 2019	CIRI-IT approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI) (Italy). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

CIRI, Italy

Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI) is a University Hospital in Genua, Italy. They participate to the DRIVE 2018/19 multi-site study with data collected using a primary care-based test-negative design (TND) study and a cohort study. This report describes the TND data
A description of the study characteristics can be obtained from the DRIVE SAP 2018/19.

The dataset used for central data quality checks was uploaded for 'analysis' using the DRIVE Electronic Study Support Application on 14/05/2019. The uploaded TND dataset contained records on 1120 patients.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks for the TND data set is provided in Table 1. Details on the data quality violations and subsequent actions taken for the TND data set are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	No, pneumovaccdate should have been pneumovaccdat. The variable has been renamed.
#3.	All records included in the dataset are ILI/SARI cases?	Yes
#4.	There are no duplicated records in the dataset?	Yes
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	Yes
#6.	ILI/SARI cases could be verified based on clinical symptoms	Yes
#7.	All records have onset date before swab date?	Yes
#8.	All records have visit date before swab date?	Yes
#9.	All records have visit date after onset date?	Yes
#10	All records have vaccination date before onset date?	Yes
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	No, 5 records violated this condition. See Table 2
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	Yes
#14	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	No, 9 records violated this condition. See Table 2
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#12	750, 1081, 1091, 1104, 1118	The variable 'labvirus' indicates 'influenza A', whereas the variable 'labsubtype' indicates 'no influenza'.	This is a coding mistake. The variable 'labsubtype' should indicate 'subtype information not available'. This is corrected as 'labsubtype1' = 9999
#14	1652, 1706, 1713, 1714, 1729, 1737, 1803, 1804, 1973	These persons are indicated to not have any chronic disease ('chronic'= 0) while the variable 'childrisk' indicates that they have perinatal or congenital risk factor.	These persons are considered having chronic disease. Correction done as follows: chronic=1

After data cleaning, a total of 1120 records were retained for further processing. After sharing the first version of this data quality report, missing information was completed by the study site upon consulting the original data collection forms (Table 3).

Table 3 Additional data quality issues, and corrections

Patient id	Data quality issue	Action
304, 594, 807, 808, 816, 1338, 2058, 2074, 2093, 2094	Missing information for variable 'pregnancy'	Changes were made upon consulting the original data collection forms: 'pregnancy' = 0
337, 1583, 1686, 2066	Missing information for variable 'seasvaccn1'	Changes were made upon consulting the original data collection forms: 'seasvaccn1' = 0
396, 2051	Missing information for variable 'seasvaccn1'	Changes were made upon consulting the original data collection forms: 'seasvaccn1' = 1
401	Mistake in visit date/swab date	Changes were made upon consulting the original data collection forms: 'swabdate' = 2018-11-05 'visitdate' = 2018-11-05
1720	Vaccination date was missing	Changes were made upon consulting the original data collection forms: 'seasvaccdate' = 2018-11-06

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. has symptom onset outside the study period
2. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
3. is less than 6 months of age at the time of the onset of the symptoms
4. has a contraindication for influenza vaccine
5. is institutionalised at the time of symptoms onset
6. will have the respiratory specimen taken ≥ 8 days after ILI onset
7. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

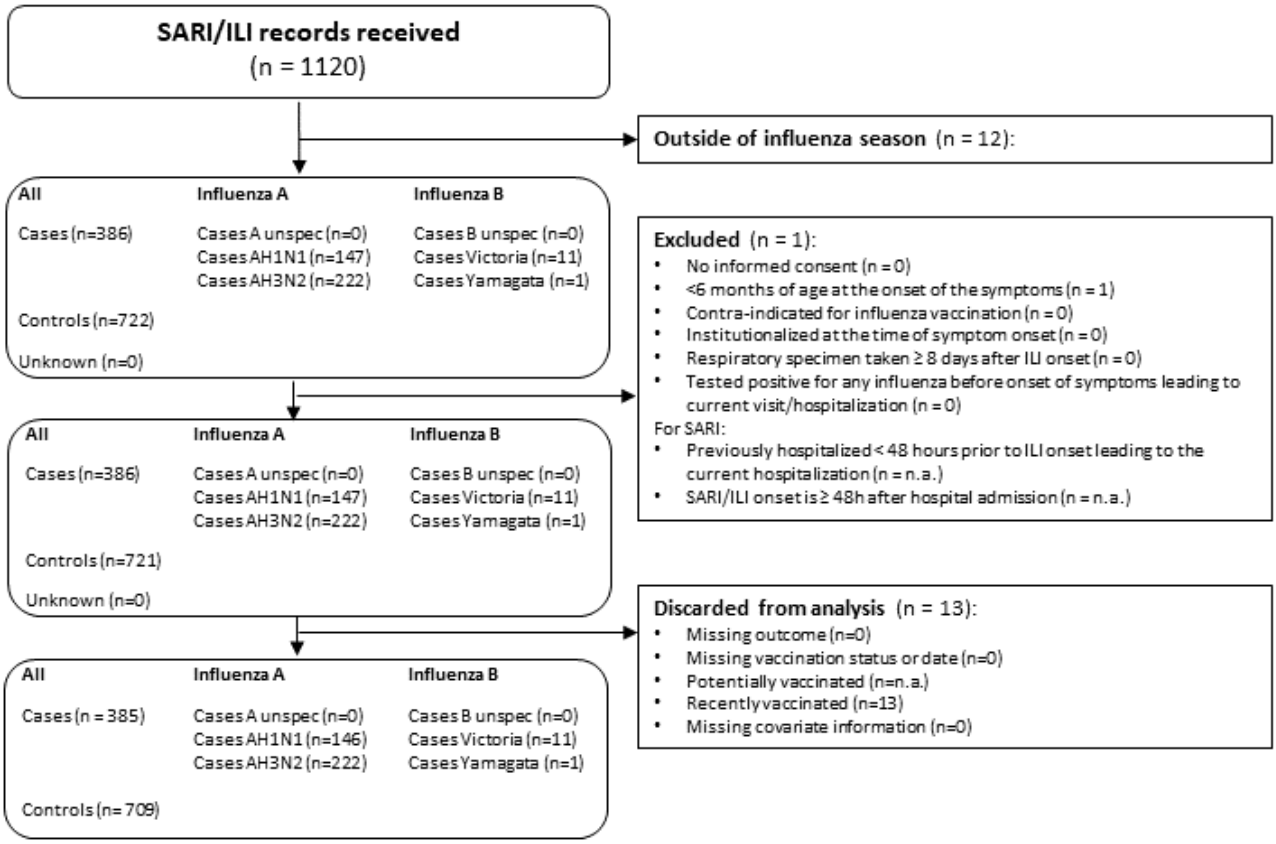


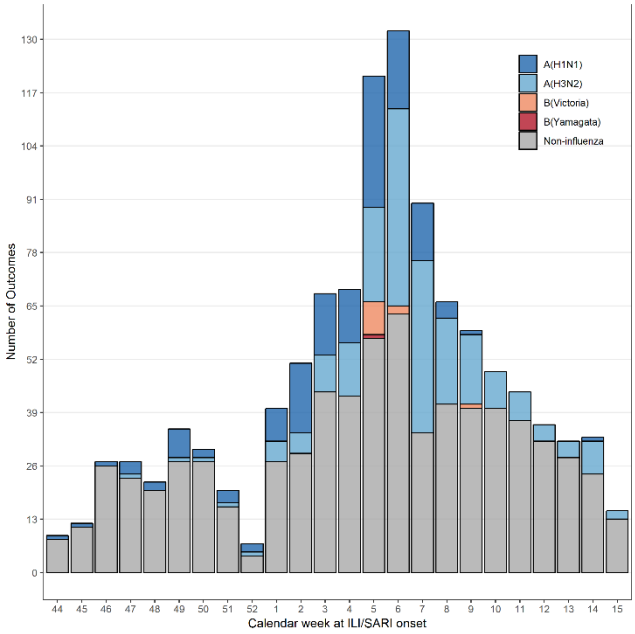
Figure 1 Attrition diagram

Data summary

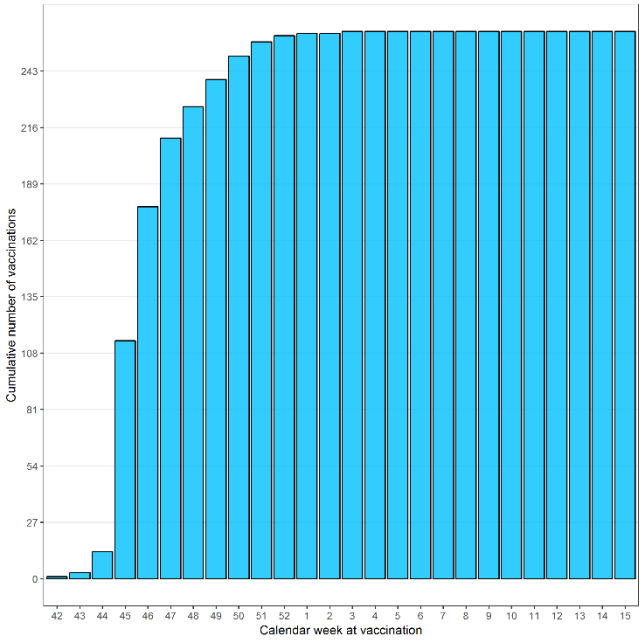
A total of 1094 subjects were retained for analysis, with the majority of subjects being either children (<17 yrs) or adults (18-64yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

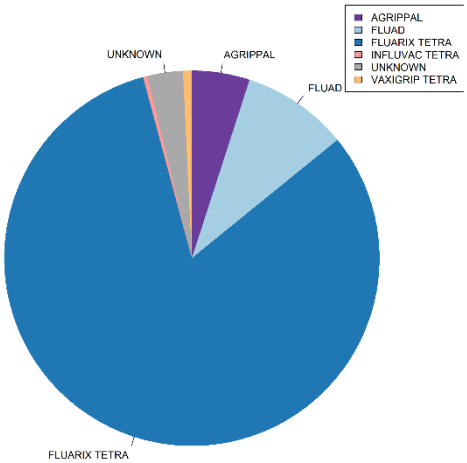
Characteristic	Number of subjects
Subjects	1094
Influenza cases	385
Vaccinated subjects	262
Age	
Children (0-17y)	384
Adults (18-64y)	520
Elderly (65+y)	190
Influenza vaccine brand	
Agripal	13
Fluad	24
Fluarix Tetra	214
Influvac Tetra	1
Vaxigrip Tetra	2
Unknown	8



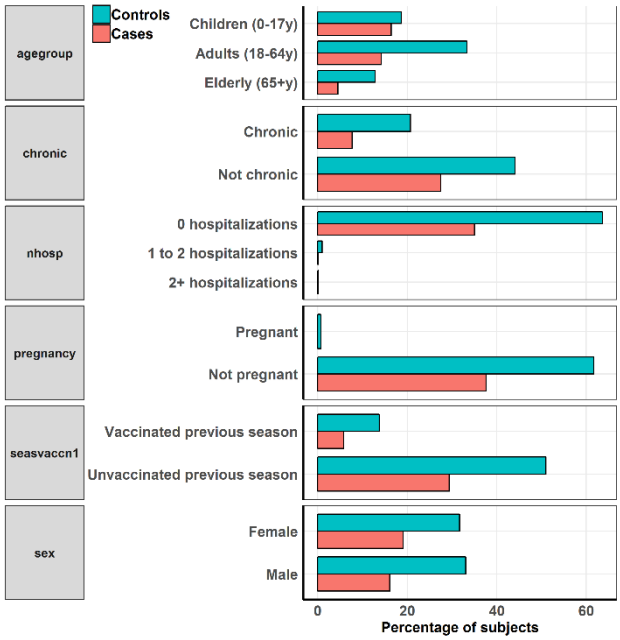
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: RCGP RSC

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Kaatje Bollaerts (P95) Nick De Smedt (P95)
Data provider	Simon de Lusignan, Chris McGee, Julian Sherlock Mana Tripathy Ivelina Yonova Harshana Liyanage Uy Hoang

Version	Date	Description
V0.1	27 May 2019	First Draft
V0.2	31 May 2019	Surrey / RCGP RSC approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) UK. The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

RCGP RSC, UK

Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) has a long established a programme of influenza and respiratory disease surveillance since 1967, making it the longest established primary care sentinel network in Europe. The RCGP RSC operates a Workload Observatory for NHS England which provides insights into current workloads and complexity of cases seen in general practice. They are actively participating in improving primary care data quality through a range of online dashboards that provide data quality metrics to participating practices. RCGP RSC contributed to the DRIVE 2018/19 multi-site study with data collected using a test-negative design study. They provided one dataset with data collected within the primary care setting. A description of the study characteristics can be obtained from the DRIVE SAP 2018/19.

The dataset used for central data quality checks was the dataset uploaded for ‘analysis’ using the DRIVE Electronic Study Support Application on 22/05/2019. The uploaded dataset contained records on 109123 patients from 258 RCGP RSC practices, of which only 6 practices participated to this study providing results on 267 influenza point-of-care tests.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check		Outcome
Nr.	**Description**	
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	No, the variable 'gpvisit' was not provided.
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	No, the variable 'vaccsource' and 'consultationcountprevyr' were additionally provided. The variable 'vaccsource' is not used in the analysis. The variable 'consultationcountprevyr' was renamed (see Table 2).
#3.	All records included in the dataset are ILI/SARI cases?	No, 126 records were indicated as ILI-negative. For these records, no onset date was provided. (See Table 2)
#4.	There are no duplicated records in the dataset?	Yes
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	No, 1 records has an outlying or invalid value for BMI. See Table 2.
#6.	ILI/SARI cases could be verified based on clinical symptoms	Additional information on all required ILI clinical symptoms was not provided. This check is not possible and this will be considered in the interpretation of the results.
#7.	All records have onset date before swab date?	Yes
#8.	All records have visit date before swab date?	Yes
#9.	All records have visit date after onset date?	Yes
#10	All records have vaccination date before onset date?	Yes
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	Yes
#14	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	Yes
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#1-2		No, the variable 'gpvisit' was not provided. The variable 'consultationcountprevyr' was.	The variable 'consultationcountprevyr' was renamed as 'gpvisit'.
#3		126 records were not indicated as ILI positive and were not having a symptom onset date	These records cannot be used in the analysis and were excluded.
#5	92906	Subjects with a very high BMI. They were all correctly indicated as suffering from 'at least one chronic condition'.	Data are correct and no changes were made.

After data cleaning, a total of 141 records were retained for further processing.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. has symptom onset outside the study period
2. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
3. is less than 6 months of age at the time of the onset of the symptoms
4. has a contraindication for influenza vaccine
5. is institutionalised at the time of symptoms onset
6. will have the respiratory specimen taken ≥ 8 days after ILI onset
7. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

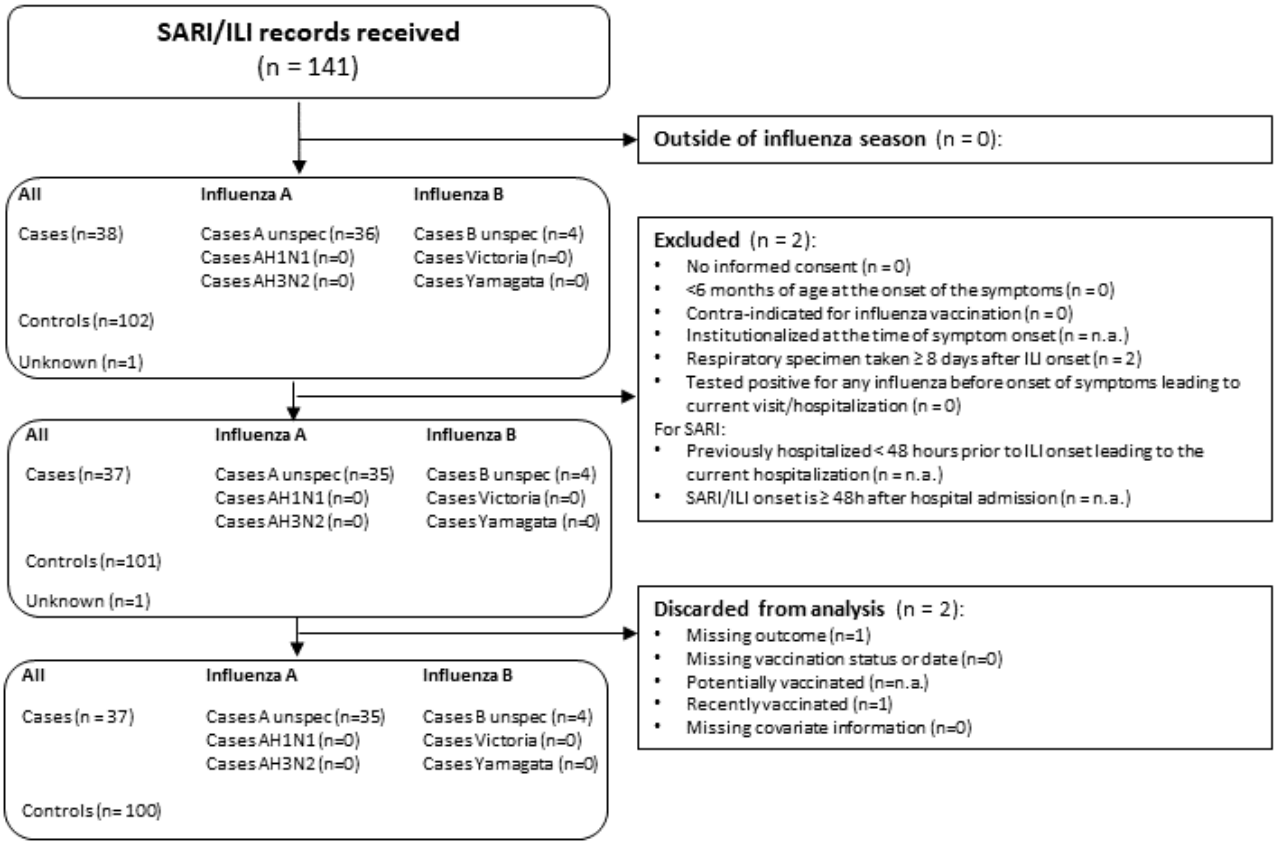


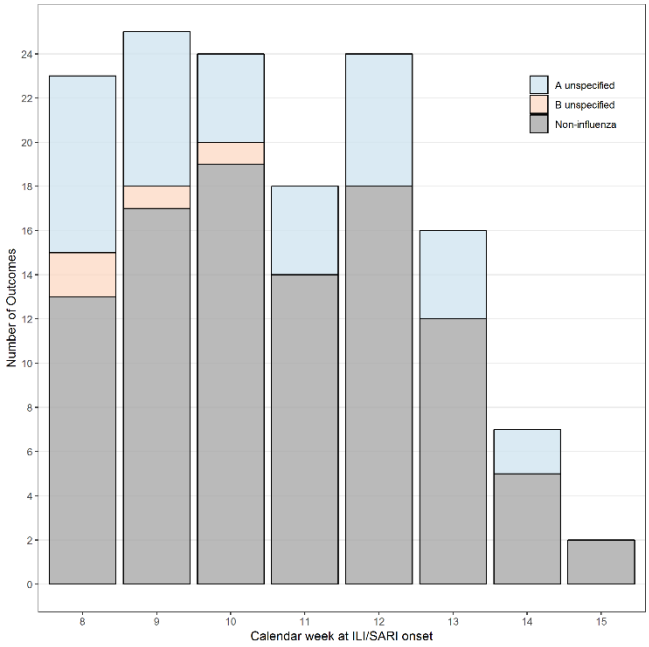
Figure 1 Attrition diagram

Data summary

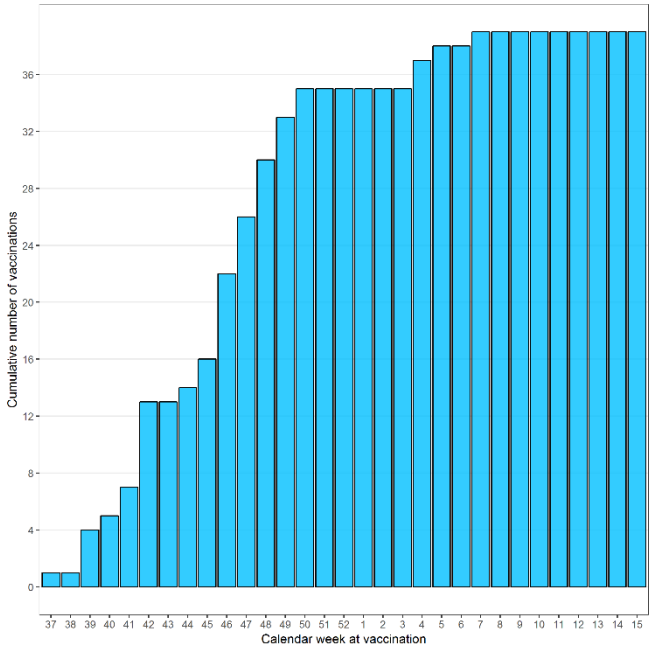
A total of 137 subjects were retained for analysis, with the majority of subjects being either children (<17 yrs) or adults (18-64yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

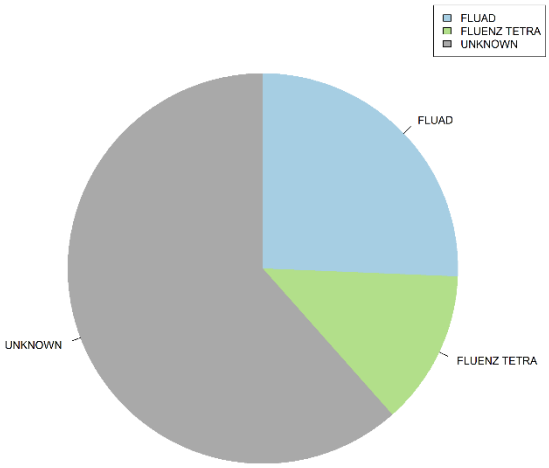
Characteristic	Number of subjects
Subjects	137
Influenza cases	37
Vaccinated subjects	39
Age	
Children (0-17y)	45
Adults (18-64y)	72
Elderly (65+y)	20
Influenza vaccine brand	
Fluad	10
Fluenz Tetra	5
Unknown	24



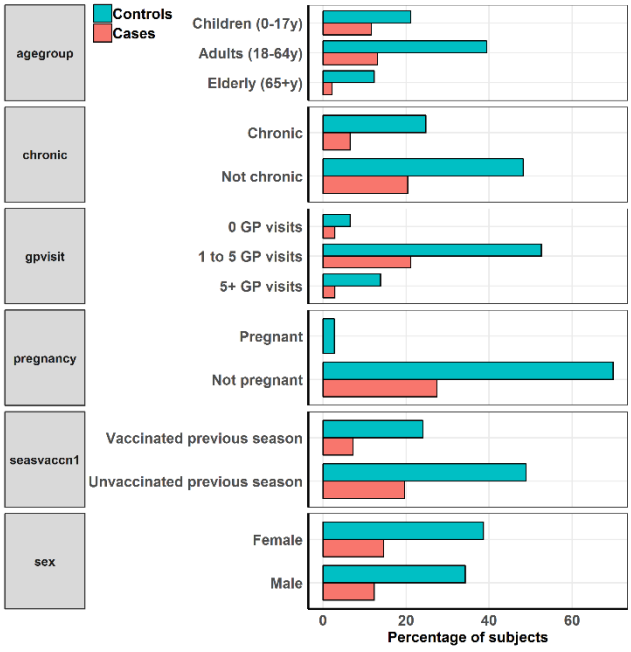
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: Istituto Superiore di Sanita, ITALY

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Maria Alexandridou (P95) Kaatje Bollaerts (P95) Nick De Smedt (P95)
Data provider	Maria Rita Castrucci, Antonino Bella

Version	Date	Description
V0.1	27 May 2019	First Draft
V1.0	05 June 2019	ISS approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Istituto Superiore di Sanita (ISS) (Italy). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

ISS, Italy

ISS is National Influenza Center at the Department of Infectious Diseases in Italy. They participate to the DRIVE 2018/19 multi-site study with data collected using a test-negative design study. They provided one dataset with data collected within the primary care setting. A description of the study characteristics can be obtained from the DRIVE SAP 2018/19.

The dataset used for central data quality checks was the dataset uploaded for ‘analysis’ using the DRIVE Electronic Study Support Application on 16/05/2019. The uploaded dataset contained records on 2573 patients.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	No, seasvaccbrand1, seasvaccbrand2, seasvaccdate1, seasvaccdate2 are not provided.
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	Yes
#3.	All records included in the dataset are ILI/SARI cases?	Yes
#4.	There are no duplicated records in the dataset?	Yes
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	Yes
#6.	ILI/SARI cases could be verified based on clinical symptoms	Additional information on ILI clinical symptoms not provided. This check is not possible and this will be considered in the interpretation of the results.
#7.	All records have onset date before swab date?	Yes
#8.	All records have visit date before swab date?	No, 1 record violated this condition. See Table 2
#9.	All records have visit date after onset date?	Yes
#10	All records have vaccination date before onset date?	No, 5 records violated this condition. See Table 2
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	No co-infections reported for this data.
#14	Variable ‘at least one chronic condition’ consistent with information provided on individual chronic conditions (when provided)?	Additional information on separate chronic conditions not provided. This check is not possible and this will be considered in the interpretation of the results.
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#8	34544	The date of swab is before GP visit date	It is only 1-day difference. No action taken
#10	16647, 18014, 18522, 18690, 23139	These persons were vaccinated after the date of ILI onset.	These persons were unvaccinated when the ILI onset occurred, hence the following changes are implemented (default values for unvaccinated persons): seasvaccany=0 seasvaccbrand=”” seasvaccdate=NA

After data cleaning, a total of 2573 records were retained for further processing.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. has symptom onset outside the study period
2. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
3. is less than 6 months of age at the time of the onset of the symptoms
4. has a contraindication for influenza vaccine
5. is institutionalised at the time of symptoms onset
6. will have the respiratory specimen taken ≥ 8 days after ILI onset
7. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

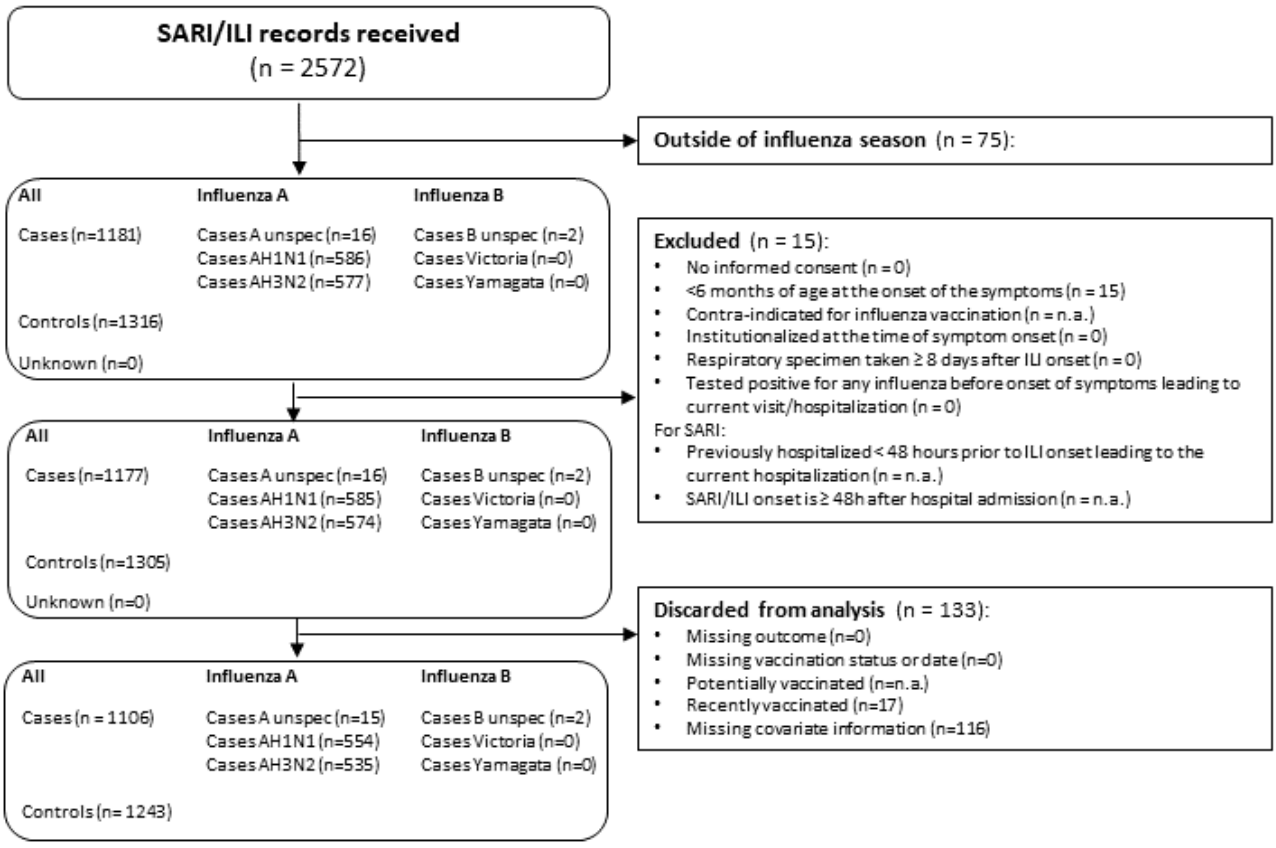


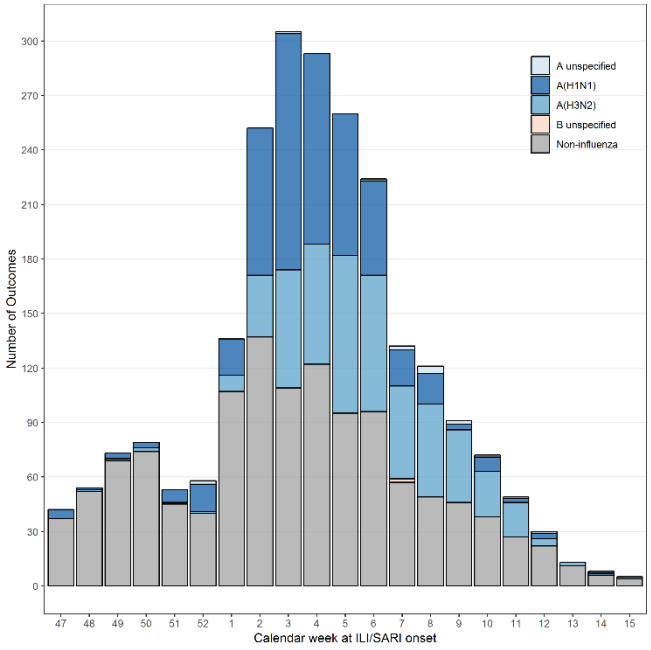
Figure 1 Attrition diagram

Data summary

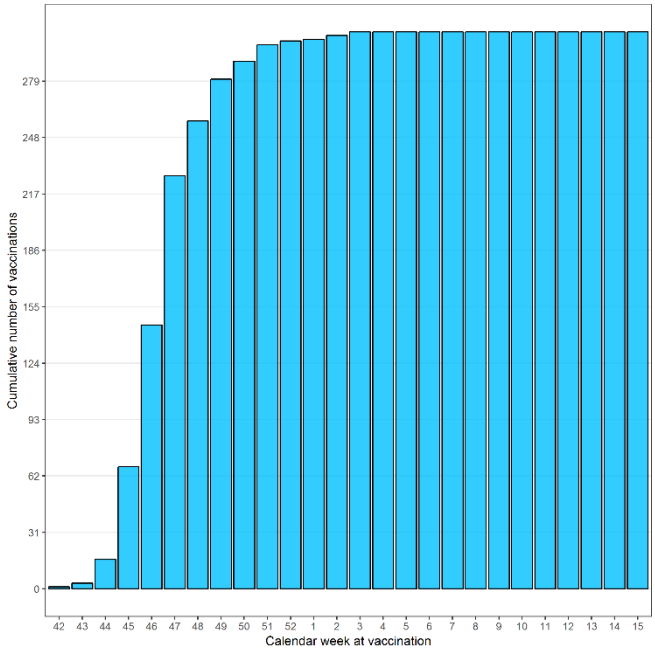
A total of 2350 subjects were retained for analysis, with the majority of subjects being either children (<17 yrs) or adults (18-64yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

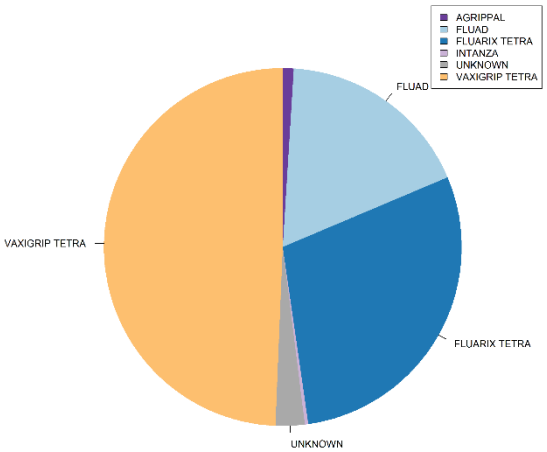
Characteristic	Number of subjects
Subjects	2350
Influenza cases	1107
Vaccinated subjects	306
Age	
Children (0-17y)	1149
Adults (18-64y)	1023
Elderly (65+y)	178
Influenza vaccine brand	
Agripal	3
Fluad	54
Fluarix Tetra	89
Intanza	1
Vaxigrip Tetra	151
Unknown	8



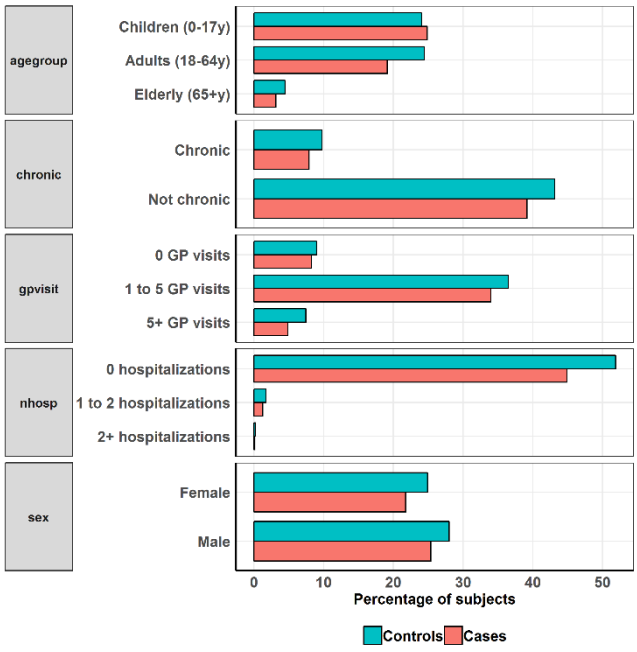
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: Helsinki University Hospital, FINLAND

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Kaatje Bollaerts (P95) Maria Alexandridou (P95) Nick De Smedt (P95)
Data provider	Raija Auvinen Ritva Syrjänen

Version	Date	Description
V0.1	27 May 2019	First Draft
V1.0	05 June 2019	Finland HUS approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Helsinki University Hospital (HUS), Finland. The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

Helsinki University Hospital, Finland

The Helsinki University Hospital (HUS) closely collaborates with THL for this study. They participate to the DRIVE 2018/19 multi-site study with data collected using a test-negative design study. They provided one dataset with data collected in hospitals. A description of the study characteristics can be obtained from the DRIVE SAP 2018/19.

The dataset used for central data quality checks were the hospital datasets uploaded for ‘analysis’ using the DRIVE Electronic Study Support Application on 17/05/2019 (revised version). The uploaded Finland hospital dataset contained records on 293 patients.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	Yes
#3.	All records included in the dataset are ILI/SARI cases?	Yes
#4.	There are no duplicated records in the dataset?	Yes
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	No, the variable 'bmi has 1 records with possibly invalid value (outlier). See Table 2
#6.	ILI/SARI cases could be verified based on clinical symptoms	Yes
#7.	All records have onset date before swab date?	Yes
#8.	For hospitalized cases, all records have hospital admission no longer than 2 days before swab date?	No, 5 records violated this condition. See Table 2
#9.	All records have visit date after onset date?	Yes
#10	All records have vaccination date before onset date?	Yes
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	Yes
#14	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	Yes
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#5	209063	The variable 'bmi has 1 records with possibly invalid value (outlier)	This patient is indeed obese. Her BMI is 55.8 which is not that extreme. Hence no action taken. The variable "obesity" also indicates that she is obese.
#8	(208865, 209713, 210729, 209795, 208843)	All records have hospital admission no longer than 2 days before swab date	Date of swab is few days before admission date. This is considered normal for hospital data.

Table 3 Additional data quality issues, and corrections

Patient id	Data quality issue	Action
304, 594, 807, 808, 816, 1338, 2058, 2074, 2093, 2094	Missing information for variable 'pregnancy'	Changes were made upon consulting the original data collection forms: 'pregnancy' = 0
337, 1583, 1686, 2066	Missing information for variable 'seasvaccn1'	Changes were made upon consulting the original data collection forms: 'seasvaccn1' = 0
396, 2051	Missing information for variable 'seasvaccn1'	Changes were made upon consulting the original data collection forms: 'seasvaccn1' = 1
401	Mistake in visit date/swab date	Changes were made upon consulting the original data collection forms: 'swabdate' = 2018-11-05 'visitdate' = 2018-11-05
1720	Vaccination date was missing	Changes were made upon consulting the original data collection forms: 'seasvaccdate' = 2018-11-06 '

We additionally found one record with vaccine brand “Vaxigrip erä R3J482v”. Erä stands for “lot” in English, therefore we identified the vaccine brand as “Vaxigrip tetra” (Table 3). After sharing the first version of this data quality report, some data information was corrected by the study site upon consulting the original data collection forms (Table 3).

Table 3 Additional data quality issues, and corrections

Patient id	Data quality issue	Action
210926	Variable ‘seasvaccbrand = ‘Vaxigrip erä R3J482v’.	The word erä was translated and identified as the English word “lot”: ‘seasvaccbrand’= VAXIGRIP TETRA
20971-3	Data entry mistake found upon checking results with swab date before visit date,	Changed from “visitdate” = 2018-02-19 to: “visitdate” = 2018-02-18
20884-3	Data entry mistake found upon checking results with swab date before visit date,	Changed from “visitdate” = 2018-03-13 to: “visitdate” = 2018-03-12
210388	Information on number of hospitalizations was missing	changed from “nhosp” = 9999 to: “nhosp” = 2
210937	Information on pregnancy was missing	changed from “pregnancy” = 9999 to: “pregnancy” = 0

After data cleaning, a total of 293 records were retained for further processing.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. is institutionalised at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after SARI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation
7. was previously hospitalised < 48 hours prior to SARI onset
8. had his/her ILI/SARI onset ≥ 48 hours after hospital admission

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

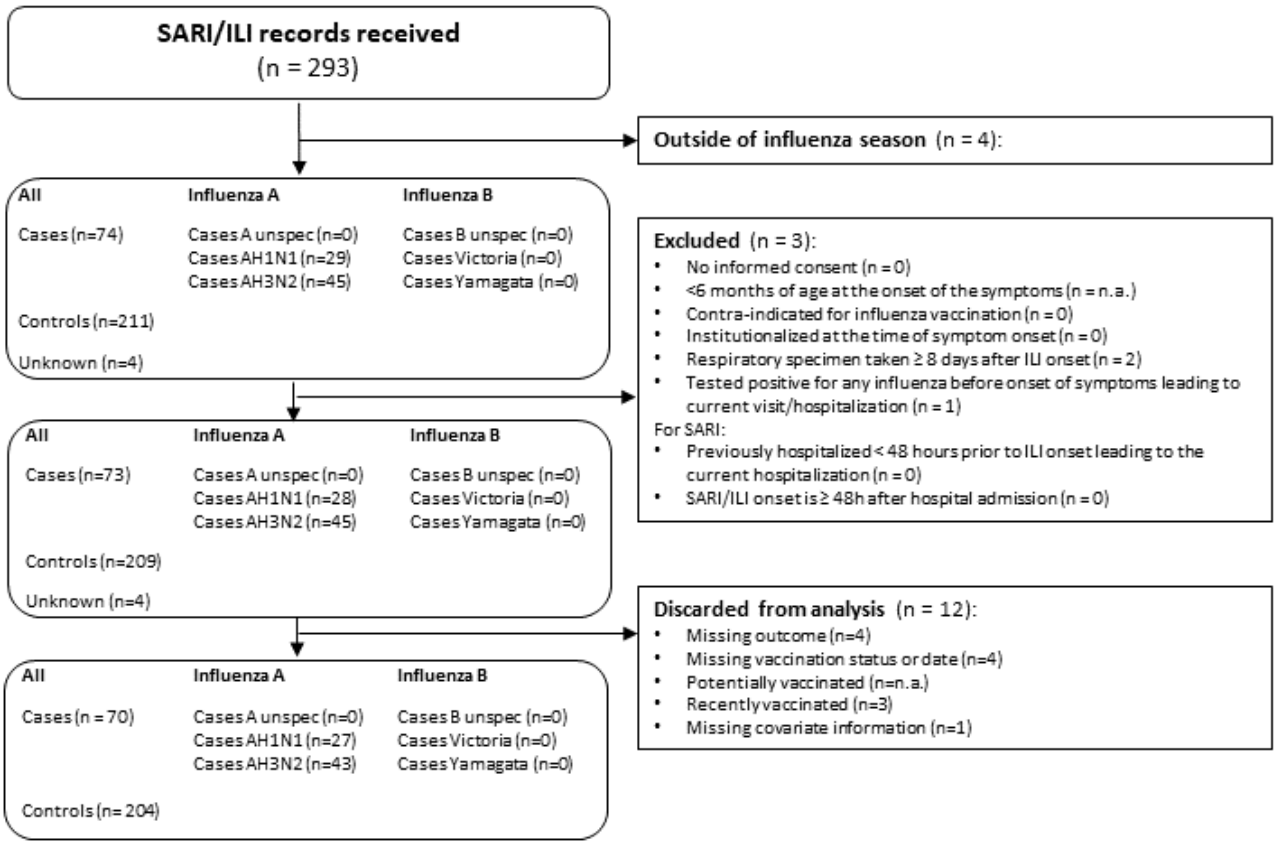


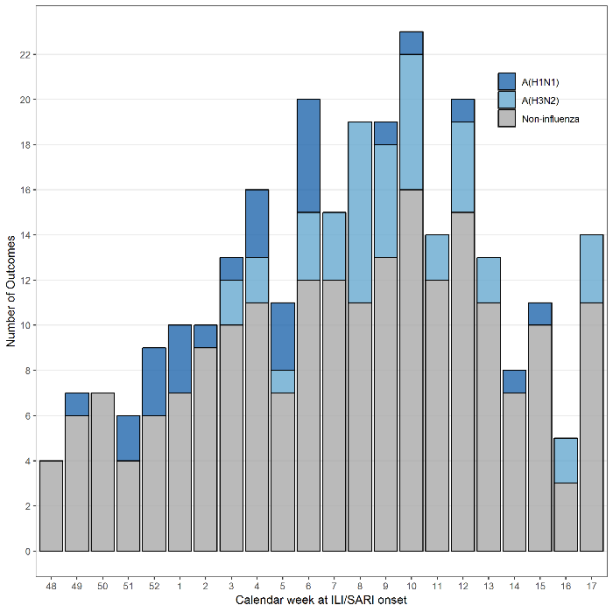
Figure 1 Attrition diagram

Data summary

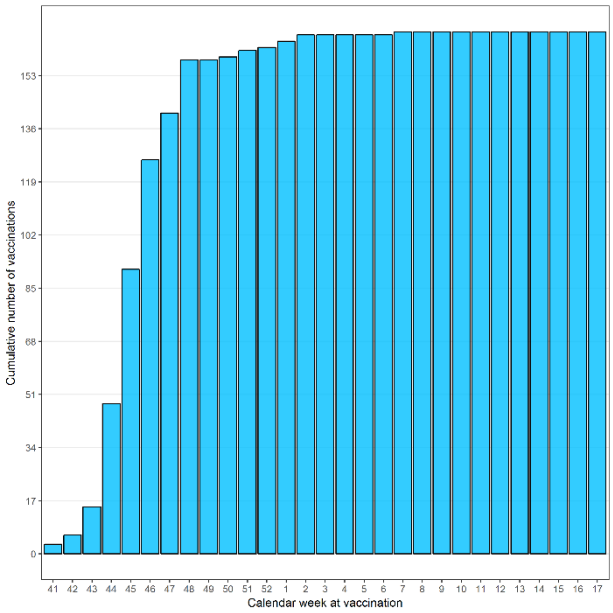
A total of 274 subjects were retained for analysis, with the majority of subjects being elderly (64+yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

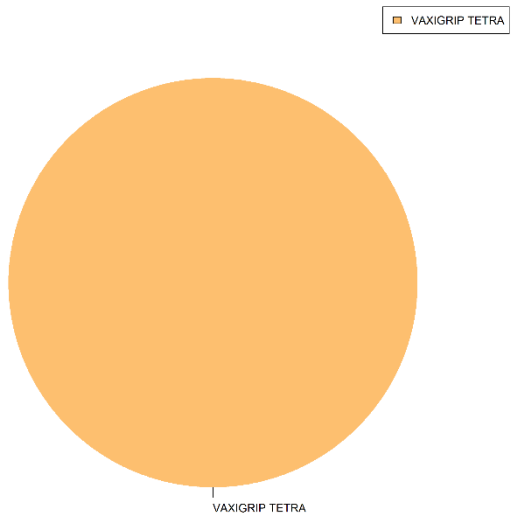
Characteristic	Number of subjects
Subjects	274
Influenza cases	70
Vaccinated subjects	167
Age	
Children (0-17y)	0
Adults (18-64y)	103
Elderly (65+y)	171
Influenza vaccine brand	
Vaxigrip Tetra	167



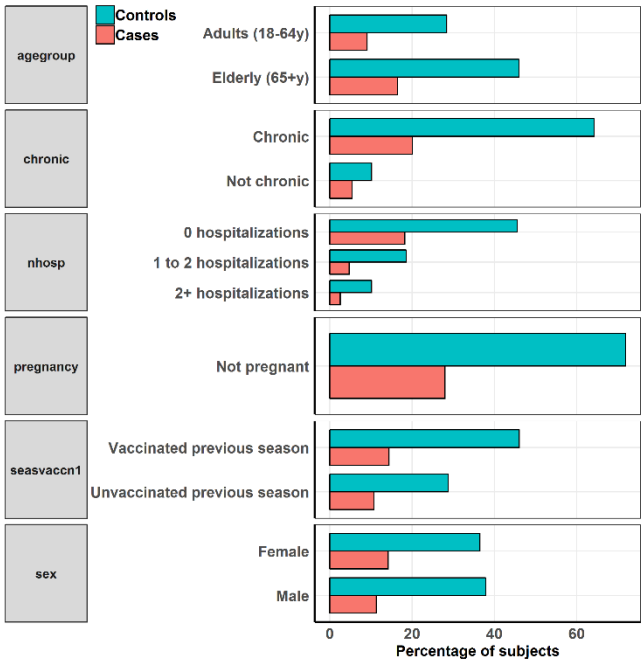
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: Hospital Network, Italy

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Kaatje Bollaerts (P95) Maria Alexandridou (P95) Nick De Smedt (P95)
Data provider	Caterina Rizzo

Version	Date	Description
V0.1	27 May 2019	First Draft
V1.0	5 June 2019	BIVE approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Hospital Network of four Italian hospitals (based in Bari, Rome, Siena and Genova). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

BIVE Hospital Network

The Italian hospital network is conducting TND hospital studies since the 2015/16 season under the I-MOVE+ project, coordinated by the Italian National Public Health Institute (Instituto Superiore di Sanità). They participate to the DRIVE 2018/19 multi-site study with data collected using a hospital-based test-negative design study.

The data sets used for central data quality checks were uploaded for ‘analysis’ using the DRIVE Electronic Study Support Application on 15/05/2019. The uploaded dataset contained records on 1675 patients.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	Yes
#3.	All records included in the dataset are ILI/SARI cases?	Yes
#4.	There are no duplicated records in the dataset?	No, 6 records contained the same information but have different patient identifiers. See Table 2.
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	No, there were 18.3% invalid values for 'smoking', 2.9% for 'seasvaccany', 13.3% for 'seasvaccn1', 25.4% for 'seasvaccn2' and 2.5% for 'deterioration'.
#6.	ILI/SARI cases could be verified based on clinical symptoms	No, the ILI/SARI case definition could not be verified for 31 (1.9%) of the cases. This will be taken into account when interpreting the results.
#7.	All records have onset date before swab date?	No, 4 records violate this condition. See Table 2.
#8.	For hospitalized cases, all records have hospital admission no longer than 2 days before swab date?	No, 8 records violate this condition. See Table 2.
#9.	All records have visit date after onset date?	Yes
#10	All records have vaccination date before onset date?	Yes
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	Yes
#14	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	Yes
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#4	(2554,2551)	Patient ids appeared to be duplicates. Checked with data site.	These are not duplicates. The data site was contacted and confirmed that both records refer to different persons with different names/dates of birth. Both records are retained in the data
#4	(2678, 2681)	Patient ids appeared to be duplicates. Checked with data site.	Same as above. Both records are retained in the data
#4	(2591, 2599)	Patient ids appeared to be	Same as above. Both records are retained

		duplicates. Checked with data site.	in the data
#5		According to the minimum data requirements, the acceptable values are 0=No, 1=Yes, 9999=No information. For all these variables, the value '8' or 'NA' was often provided.	The data site was contacted and it was confirmed that the values '8' or 'NA' indicate 'no information'. The values of these variables were changed accordingly to satisfy the minimum data requirements.
#7	2084	Date of swab is before visit date and onset date	Upon contacting the site, the swab date was corrected to 12/12/2018
#7	2927	Date of swab is before visit date and onset date	Upon contacting the site, the swab date was corrected to 26/03/2018
#7	2961	Date of swab is before visit date and onset date	Upon contacting the site, the swab date was corrected to 04/03/2018
#7	2142	Date of swab is before visit date and onset date	Upon contacting the site, the swab date was corrected to 20/02/2018
#8	2084	Date of swab is before visit date	Was corrected above (#7)
#8	1835	Date of swab is before visit date	Upon contacting the site, the swab date was corrected to 07/01/2019
#8	1475	Date of swab is before visit date	Upon contacting the site, the swab date was corrected to 16/01/2019
#8	1469	Date of swab is before visit date	Upon contacting the site, the swab date was corrected to 19/01/2019
#8	2927	Date of swab is before visit date	Was corrected above (#7)
#8	2961	Date of swab is before visit date	Was corrected above (#7)
#8	2142	Date of swab is before visit date	Was corrected above (#7)
#8	2556	Date of swab is before visit date	Upon contacting the site, the swab date was corrected to 25/03/2019

After data cleaning, a total of 1675 records were retained for further processing. We additionally found 260 records for which it was reported that no informed consent was given. The site confirmed that this is not possible and that only records with informed consent were shared.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. is institutionalised at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after SARI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation
7. was previously hospitalised < 48 hours prior to SARI onset
8. had his/her ILI/SARI onset ≥ 48 hours after hospital admission

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

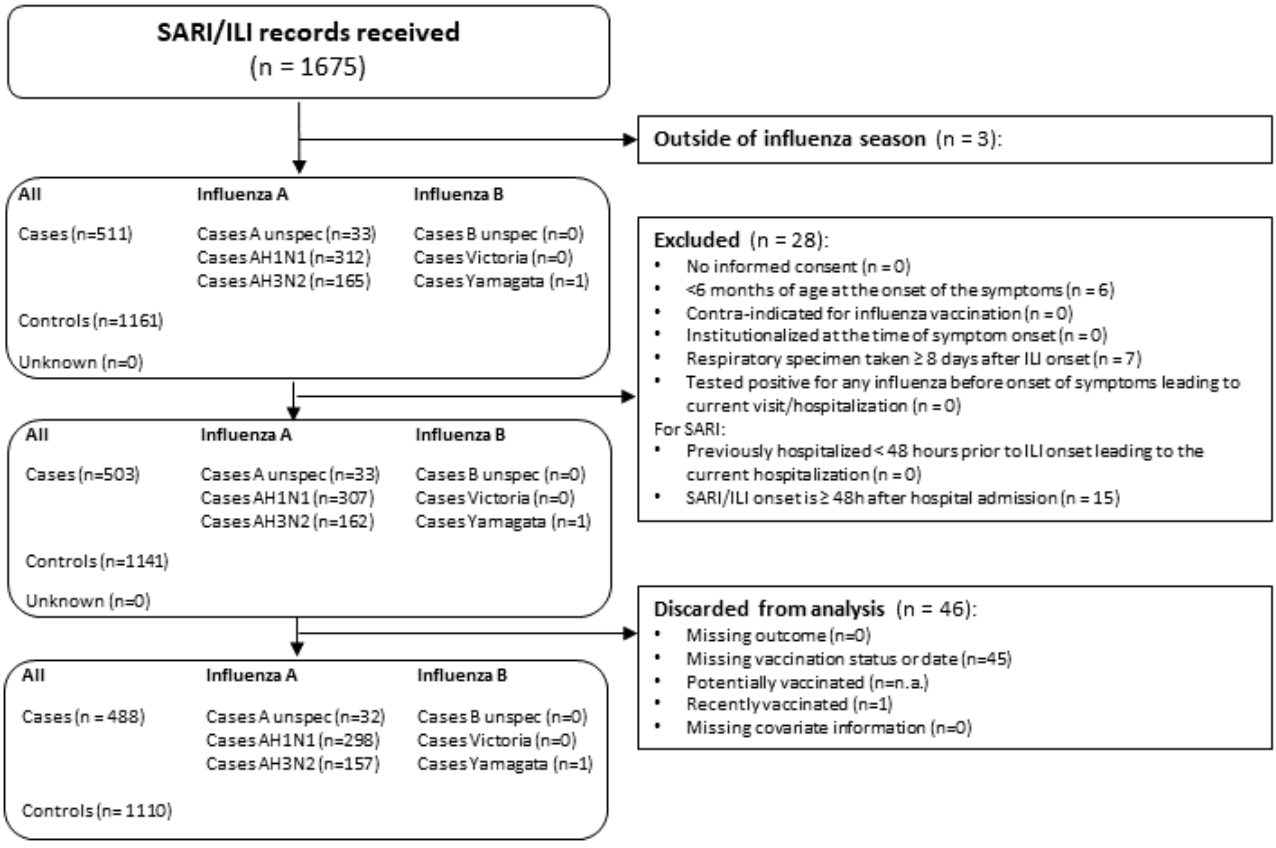


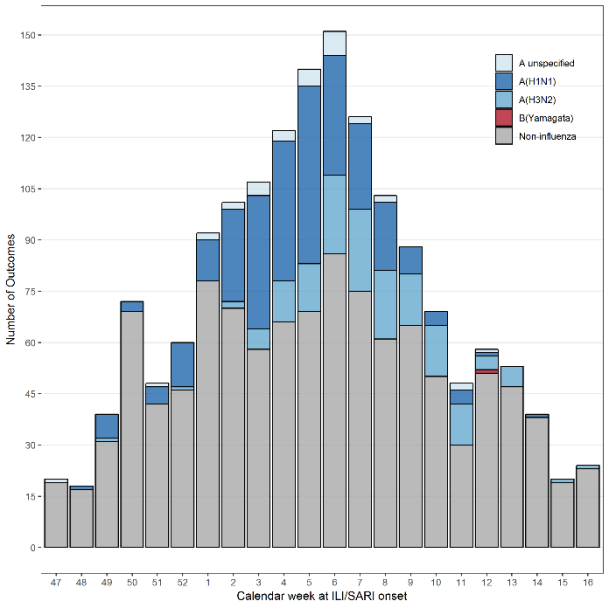
Figure 1 Attrition diagram

Data summary

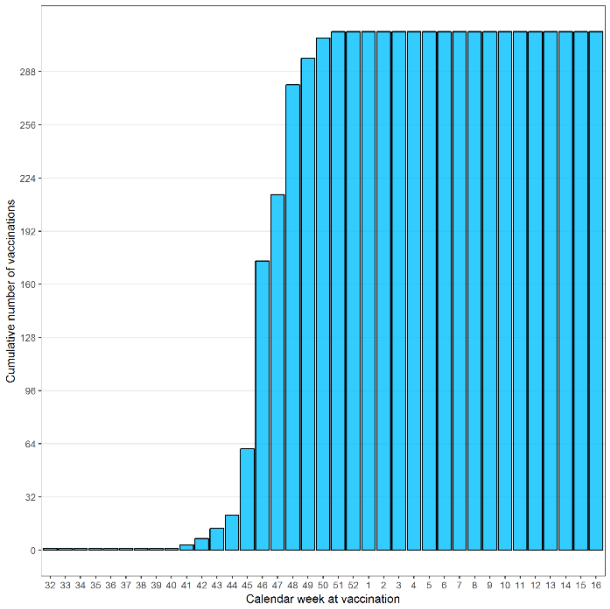
A total of 1598 subjects were retained for analysis, with the majority of subjects being children (<17 yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

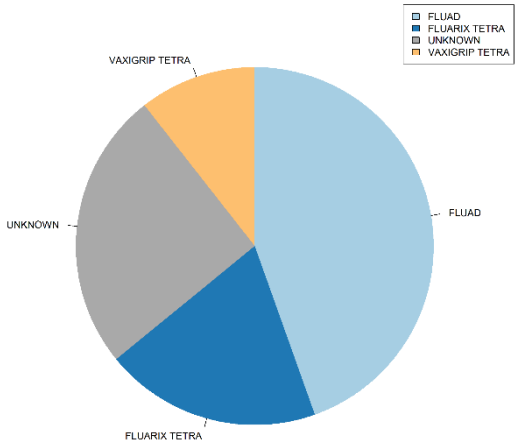
Characteristic	Number of subjects
Subjects	1598
Influenza cases	488
Vaccinated subjects	312
Age	
Children (0-17y)	820
Adults (18-64y)	278
Elderly (65+y)	500
Influenza vaccine brand	
Fluad	139
Fluarix Tetra	61
Vaxigrip Tetra	33
Unknown	79



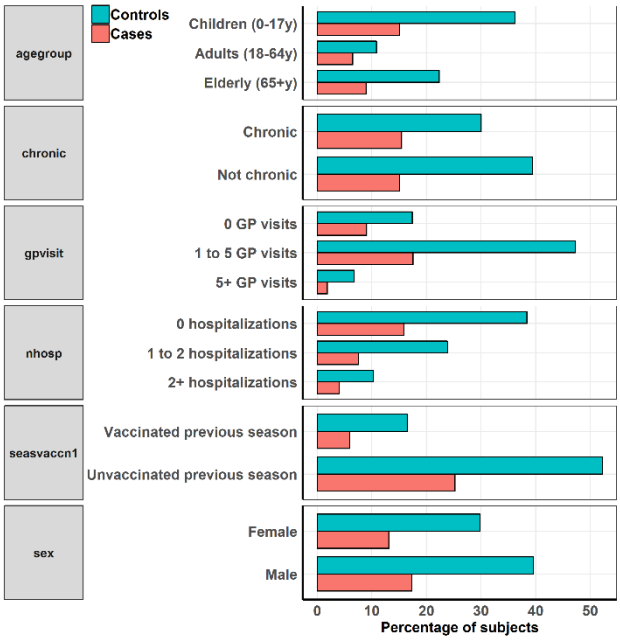
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: National Institute for Infectious Diseases "Prof. Dr. Matei Balș", Romania

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Kaatje Bollaerts (P95) Maria Alexandridou (P95) Nick De Smedt (P95)
Data provider	Oana Săndulescu,

Version	Date	Description
V0.1	27 May 2019	First Draft
V0.2	5 June 2019	Approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the National Institute for Infectious Diseases 'Prof. Dr. Matei Bals' (Romania). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

National Institute for Infectious Diseases "Prof Dr. Matei Bals"

The National Institute for Infectious Diseases "Prof. Dr. Matei Balș" is a research center and the excellence center in Romania for prevention, management and research in the area of infectious diseases and epidemiology. The Institute is a reference center for patients with influenza from Bucharest and the surrounding areas. They participate to the DRIVE 2018/19 multi-site study with data collected using a hospital-based test-negative design study.

The data sets used for central data quality checks were uploaded for 'analysis' using the DRIVE Electronic Study Support Application on 21/05/2019. The uploaded dataset contained records on 1619 patients, of which 1156 SARI patients.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	Yes
#3.	All records included in the dataset are ILI/SARI cases?	Yes
#4.	There are no duplicated records in the dataset?	Yes
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	No, the variable 'visitdate' had 3 records with invalid values. See Table 2
#6		No, the variable 'bmi' has 4 records with invalid or outlying values.
#7		No, the variable 'seasvaccdate' has 1 invalid value
#8		No, the variable 'seasvaccdate2' has 1 invalid value
#9		No, the variable 'pneumovaccdate' has 1 invalid value
#10	ILI/SARI cases could be verified based on clinical symptoms	Yes
#11.	All records have onset date before swab date?	No, 2 records violated this condition. See Table 2
#12.	For hospitalized cases, all records have hospital admission no longer than 2 days before swab date?	No, 4 records violated this condition. See Table 2
#13.	All records have visit date after onset date?	No, 16 records violated this condition. See Table 2
#14.	All records have vaccination date before onset date?	No, 1 record violated this condition. See Table 2
#15	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	No, 1 record violated this condition. See Table 2
#16	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#17	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	Yes
#18	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	No, 37 records were recorded as 'no chronic condition' whereas a chronic condition ('cardiovasc', 'anemia', 'childrisk') was present
#19	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#5	G11, G12, A76	Visit date had invalid (future) values.	Visit dates were end 2019, whereas onset dates were end 2018. These typos were corrected, and year 2019 was replaced by 2018.
#6	135, 412, G42, A21	BMI of these patients < 10, below limits of survival.	In the analysis, the variable BMI is only used to identify obesity (BMI > 30) as a chronic condition. Giving the limited

			importance of this variable, the impossible values are indicated as 'missing information' for patients 135, 412, G42. For patient A21, BMI was corrected as 'bmi' = 20.4.
#7	G178	'Seasvaccdate' has an implausible value (01/10/2019)	The value was changed to (20/11/2018)
#8	G178	'Seasvaccdate2' has an implausible value (01/10/2019)	The value was changed to (10/1/2019)
#9	239	Pneumovaccdate has invalid date value (13/07/2019)	The value was changed to (13/07/2018)
#11	263	Inconsistencies in swab date, onset date and visit date: onset date after swab date/visit date: Visitdate/onsetdate = 27/01/2018 and onsetdate = 25/01/2019.	The typo in the year of the Visitdate/onsetdate has been corrected to 27/01/2019
#11	G292	Inconsistencies in swab date and visit date	Impossible value for swabdate. Upon contacting the site, the swabdate was set equal to the visitdate (swabdate= 29/03/2019)
#12	436	Inconsistencies in swab date and visit date Onsetdate: 20/03/19 Visitdate: 27/03/19 Swabdate: 22/03/19	Visitdate was corrected as: 21/03/19
#12	G12	Inconsistencies in swab date and visit date Onsetdate: 29/12/18 Visitdate: 31/12/18 Swabdate: 03/01/19	Possible that patient got swabbed at primary care before admitted to hospital. No changes made
#12	G292	Inconsistencies in swab date and visit date	See #11 (patientID =G292) above. Corrections made above
#12	A76	Inconsistencies in swab date and visit date	See # 5 (patientID = A76) above. Corrections made above
#13	A04, A120, A200, A254, A279, A339, A355, A384, GA98	Vaccination status is unknown as these patients were not eligible for the study. No swabs were obtained.	These records will be discarded when applying the study exclusion criteria (see next Section)
#13	263		See #11 (ID = 263) above. Corrections made above.
#13	A05	Inconsistencies in swab date and visit date Onsetdate: 02/12/18 Visitdate: 12/06/18 Swabdate: 07/12/18	Month and day value are swapped. Correct visit date to 06/12/2018

#13	A06	Inconsistencies in swab date and visit date Onsetdate: 06/12/18 Visitdate: 12/10/18 Swabdate: 10/12/18	Month and day value are swapped. Correct visit date to 10/12/2018
#13	A07	Inconsistencies in swab date and visit date Onsetdate: 06/12/18 Visitdate: 12/10/18 Swabdate: 10/12/18	Month and day value are swapped. Correct visit date to 10/12/2018
#13	A08	Inconsistencies in swab date and visit date Onsetdate:11/12/2018 Visitdate: 12/11/2018 Swabdate: 13/12/2018	Month and day value are swapped. Correct visit date to 11/12/2018
#13	A09	Inconsistencies in swab date and visit date Onsetdate: 11/12/2018 Visitdate: 12/11/2018 Swabdate: 12/12/2018	Month and day value are swapped. Correct visit date to 11/12/2018
#13	GA60	Inconsistencies in swab date and visit date Onsetdate: 17/01/2019 Visitdate: 18/01/2018 Swabdate: 21/01/2019	Wrong year of visit date. Correct visit date to 18/01/2019
#14	G178	Inconsistency in vaccination date and onset date	See #7 and #8 above. Corrections made above
#15	G237	Inconsistency in vaccination status and date at vaccination	The following data has been corrected seasvaccany = 1 seasvaccdate1 = 21/01/2019 seasvaccdate2 = 21/02/2019
#18	18, 43,51,52,78,103,120,121,145,177 ,190,194,198,205,293,333,334,37 9,384,397,412,415,482,502, G32, G37, G40, G43, G59, G107, G122, G164, G222, G237, G291, G305, G332	Variable '>1 chronic condition' is inconsistent with variables 'cardiovasc', 'anemia' and 'childrisk'. ('chronic = 0' & ('cardiovasc' = 1 OR amenia = 1 OR childrisk = 1)).	The variable 'chronic' means 'at least 1 chronic condition' out of a list of chronic conditions (including obesity). All these records have been relabelled as 'chronic' = 1

After data cleaning, a total of 1156 records were retained for further processing. . After sharing the first version of this data quality report, missing information was completed by the study site upon consulting the original data collection forms (Table 3).

Table 3 Additional data quality issues, and corrections

Patient id	Data quality issue	Action
A341, A392	Missing outcome information.	The original data was checked, and confirmed that both subjects tested negative on PCR for influenza/RSV. The following change was made: 'labvirus1' = 0

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. is institutionalised at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after SARI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation
7. was previously hospitalised < 48 hours prior to SARI onset
8. had his/her ILI/SARI onset ≥ 48 hours after hospital admission

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

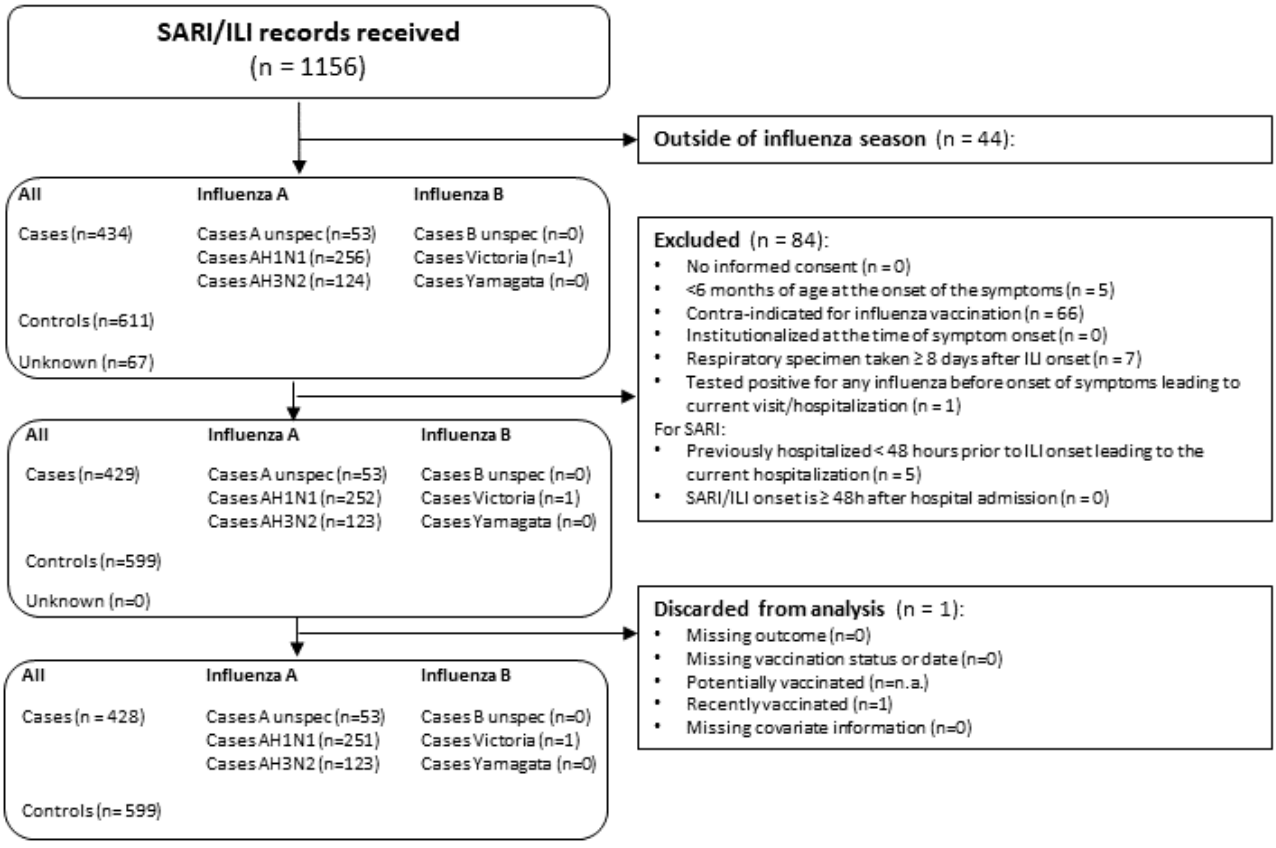


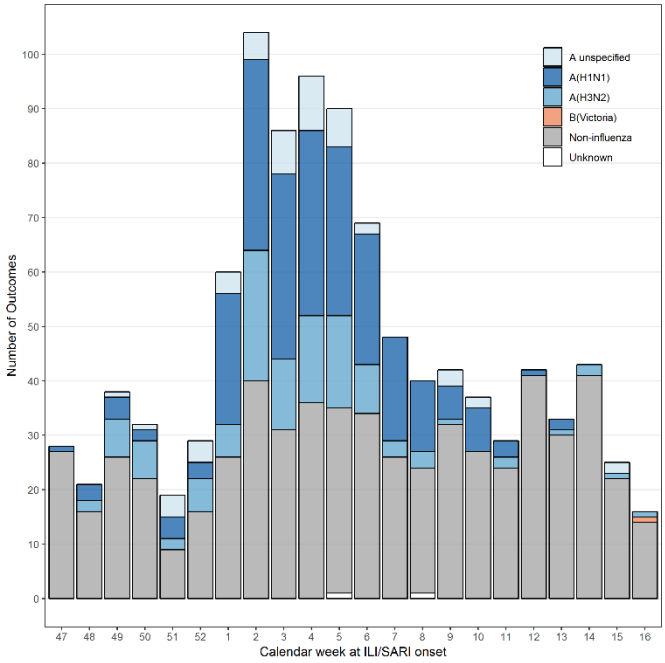
Figure 1 Attrition diagram

Data summary

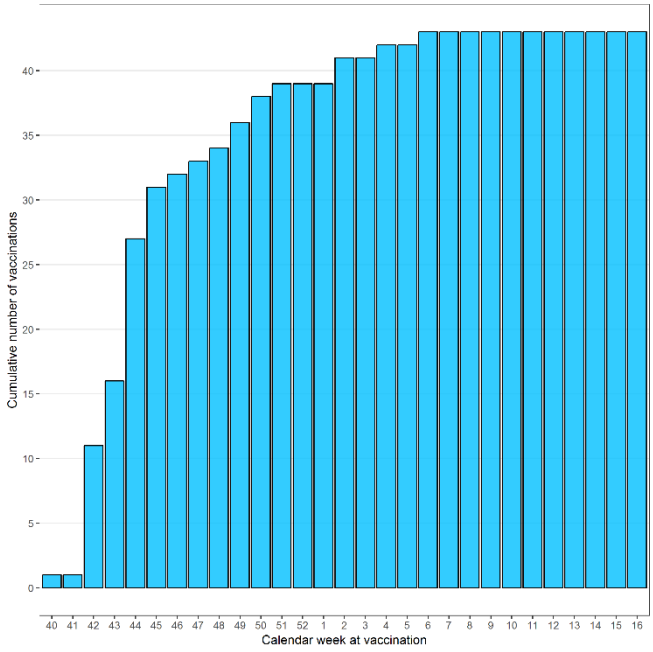
A total of 1027 subjects were retained for analysis, with the majority of subjects being either children (<17 yrs) or adults (18-64yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

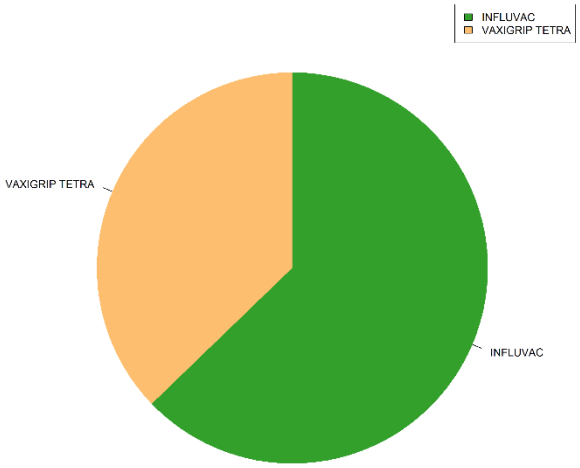
Characteristic	Number of subjects
Subjects	1027
Influenza cases	428
Vaccinated subjects	43
Age	
Children (0-17y)	518
Adults (18-64y)	356
Elderly (65+y)	153
Influenza vaccine brand	
Infuvac	27
Vaxigrip Tetra	16



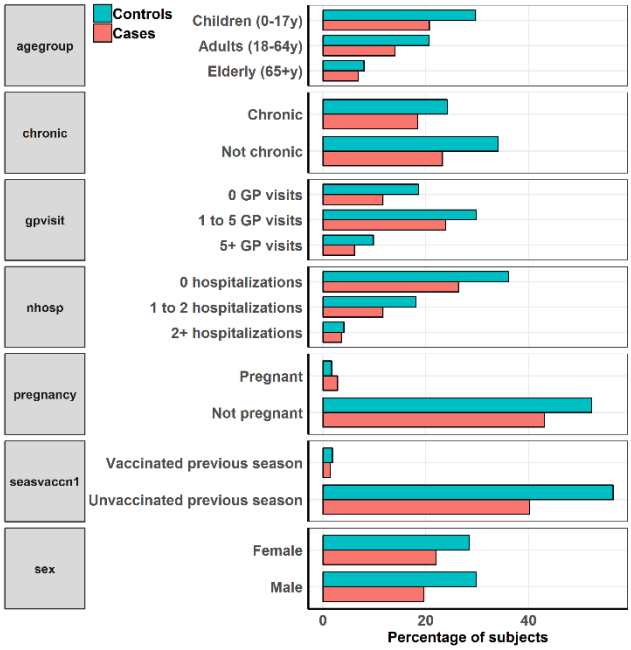
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: FISABIO, Spain

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Maria Alexandridou (P95) Kaatje Bollaerts (P95) Nick De Smedt (P95)
Data provider	Ainara Mira Iglesias

Version	Date	Description
V0.1	28 May 2019	First Draft
V1.0	5 June 2019	FISABIO approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by FISABIO (Valencia, Spain). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

FISABIO

Since 2011, FISABIO conducts an observational study applying an active annual surveillance system scheme in the Valencia Hospital Network for the Study of Influenza (VAHNSI). They participate to the DRIVE 2018/19 multi-site study with data collected using a hospital-based test-negative design study.

The data sets used for central data quality checks were uploaded for ‘analysis’ using the DRIVE Electronic Study Support Application on 16/05/2019. The uploaded dataset contained records on 7019 patients, of which 3615 are ILI/SARI patients, comprising of 2743 patients reported to be ILI/SARI, 619 patients < 5 years from whom a swab was taken and 253 patients that fulfilled the DRIVE ILI/SARI case definition but had symptom onset > 7 days before swabbing.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	Yes
#3.	All records included in the dataset are ILI/SARI cases?	No, more data than strictly needed were provided. A total of 3657 cases were removed as they did not satisfy the ILI/SARI case definition
#4.	There are no duplicated records in the dataset?	No, 2 records contained the same information but have different patient identifiers. See Table 2.
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	No, 4 records were having invalid or outlying values for bmi. See Table 2.
#6.	ILI/SARI cases could be verified based on clinical symptoms	No, the ILI/SARI case definition could not be verified for 619 (17.1%) of the cases. This will be taken into account when interpreting the results.
#7.	All records have onset date before swab date?	Yes
#8.	For hospitalized cases, all records have hospital admission no longer than 2 days before swab date?	Yes
#9.	All records have visit date after onset date?	Yes
#10	All records have vaccination date before onset date?	No, 94 (2.6%) records violated this condition. See Table 2
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	No, 58 (1.6%) records violated this condition. See Table 2
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	Yes
#14	Variable ‘at least one chronic condition’ consistent with information provided on individual chronic conditions (when provided)?	Yes
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#4	(100364,100365)	Patient ids appeared to be duplicates. Checked with data site.	These are not duplicates. The data site was contacted and confirmed that both records refer to different persons with different names/dates of birth. Both records are retained in the data
#5	102330, 20306, 21074, 21854	Subjects with a very high BMI.	Data are correct and no changes were

	They were all correctly indicated as suffering from ‘at least one chronic condition’.	made.
#10	Subjects were vaccinated after ILI/SARI onset. They are considered unvaccinated.	The following changes were made: ‘seasvaccany’ = 0 ‘seasvaccbrand’ = “ ” ‘seasvaccdate’= “NA
#12	The variable ‘labvirus’ indicates ‘influenza A’, whereas the variable ‘labsubtype’ indicates ‘no influenza’.	This is a coding mistake. The variable ‘labsubtype’ should indicate ‘subtype information not available’. This is corrected as ‘labsubtype1’ = 9999

After data cleaning, a total of 3615 records were retained for further processing.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. is institutionalised at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after SARI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation
7. was previously hospitalised < 48 hours prior to SARI onset
8. had his/her ILI/SARI onset ≥ 48 hours after hospital admission

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

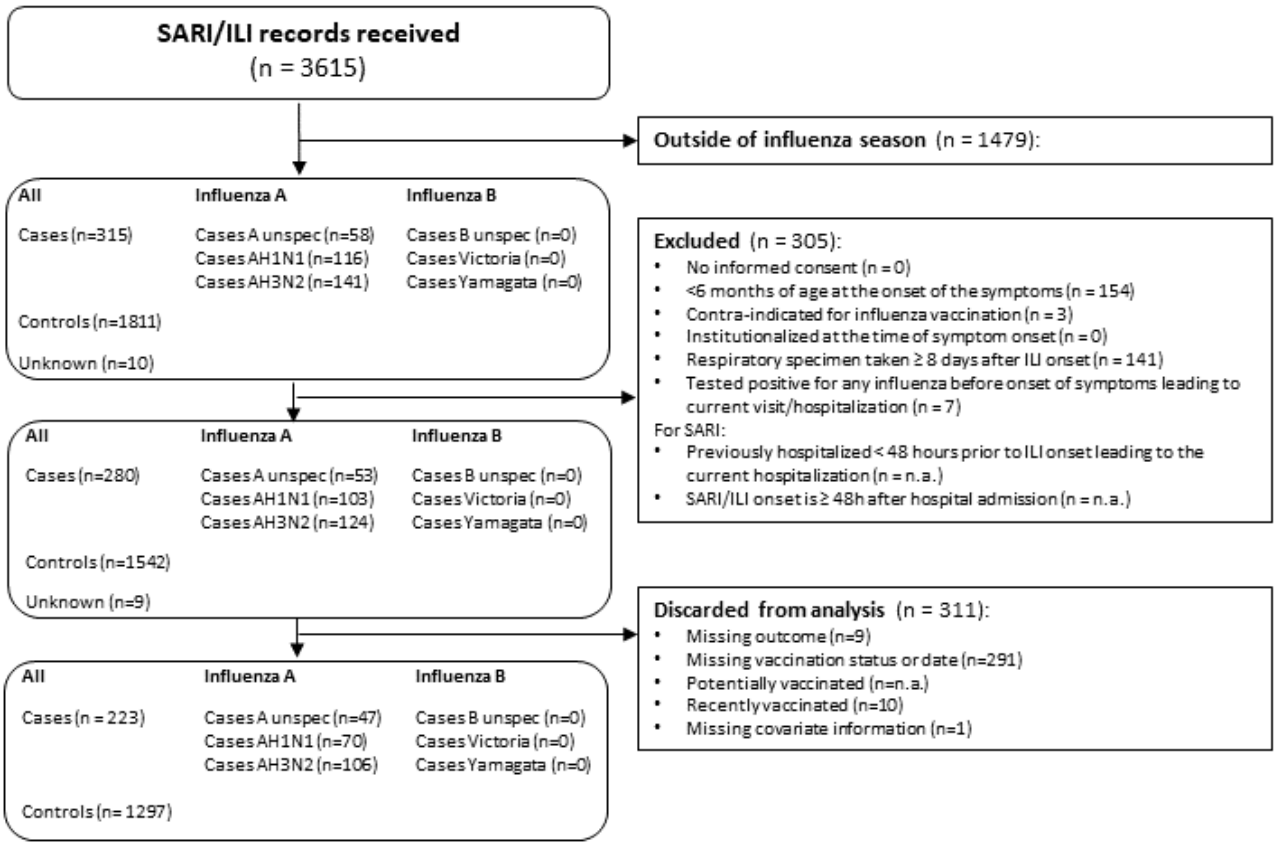


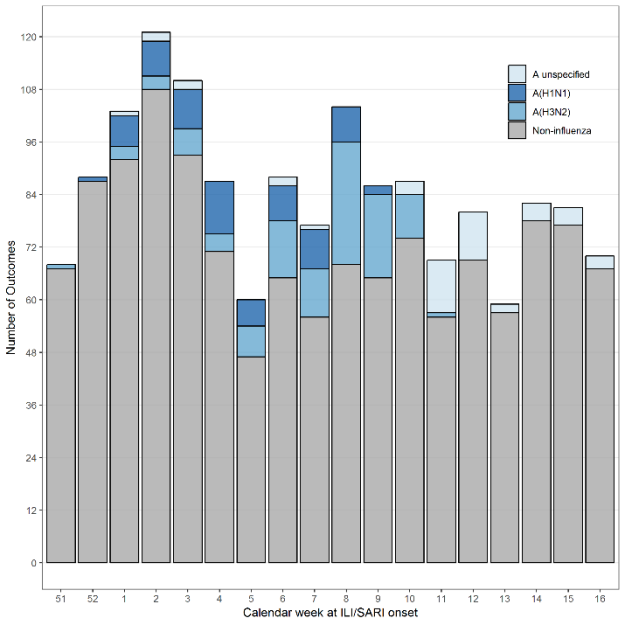
Figure 1 Attrition diagram

Data summary

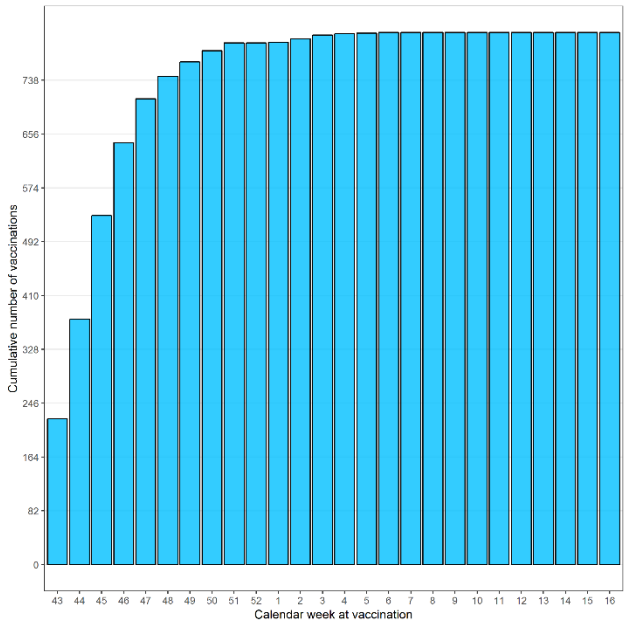
A total of 1520 subjects were retained for analysis, with the majority of subjects being elderly (65+yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

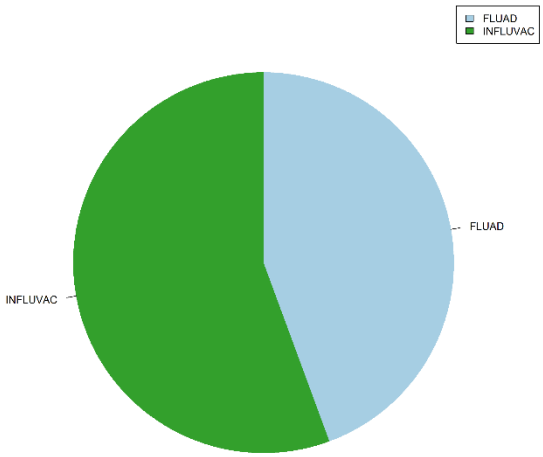
Characteristic	Number of subjects
Subjects	1520
Influenza cases	223
Vaccinated subjects	811
Age	
Children (0-17y)	187
Adults (18-64y)	234
Elderly (65+y)	1099
Influenza vaccine brand	
Fluad	360
Influvac	451



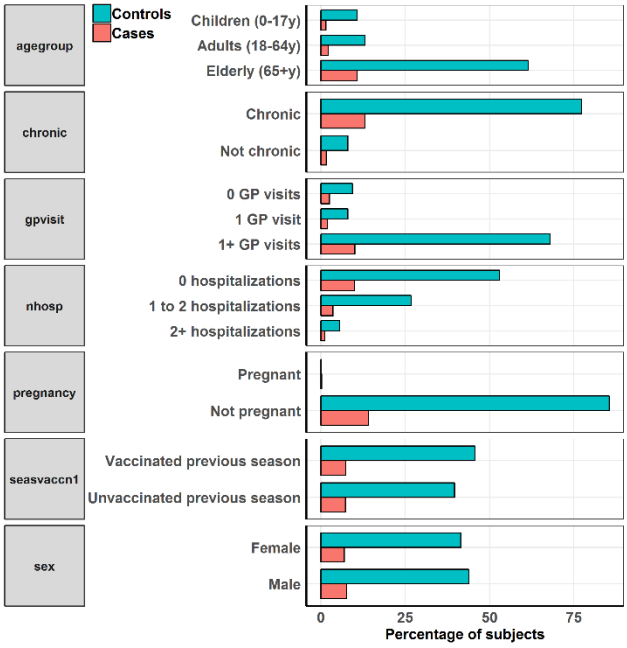
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: Vall d'Hebron, Spain

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Maria Alexandridou (P95) Kaatje Bollaerts (P95) Nick De Smedt (P95)
Data provider	Eva del Amo Moran

Version	Date	Description
V0.1	27 May 2019	First Draft
V0.2	05 June 2019	Approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the University Hospital Vall d'Hebron in Spain. The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements

Vall d'Hebron

Vall d'Hebron is a university hospital in Spain. They participate to the DRIVE 2018/19 multi-site study with data collected using a hospital-based test-negative design study.

The data sets used for central data quality checks were uploaded for 'analysis' using the DRIVE Electronic Study Support Application on 15/05/2019. The uploaded dataset contained records on 468 patients.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements?	Yes
#3.	All records included in the dataset are ILI/SARI cases?	Yes
#4.	There are no duplicated records in the dataset?	Yes
#5	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements?	Yes
#6.	ILI/SARI cases could be verified based on clinical symptoms	Additional information on ILI clinical symptoms not provided. The site reported that all ILI/SARI cases were confirmed based on clinical symptoms. The variables on the symptoms were not provided as they were not obligatory.
#7.	All records have onset date before swab date?	Yes
#8.	For hospitalized cases, all records have hospital admission no longer than 2 days before swab date?	Yes
#9.	All records have visit date after onset date?	Yes
#10	All records have vaccination date before onset date?	Yes
#11	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#12	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	No, inconsistency found for 13 records. See Table 2.
#13	Information on influenza subtype/lineage consistent with information on influenza type (co-infection)?	No co-infections reported for this data.
#14	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	Additional information on separate chronic conditions not provided. The site reported that all subjects with at least 1 chronic condition could be confirmed based on individual chronic conditions. The variables on the individual chronic conditions were not provided as they were not obligatory.
#15	All pregnant persons are women?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#12	18836107, 11907679, 11947240, 18096038, 11764762, 14859083, 11418063, 12045994, 18220577, 11618949, 15016949, 11663568, 19120837	The variable 'labvirus' indicates 'influenza A', whereas the variable 'labsubtype' indicates 'no influenza'.	This is a coding mistake. The variable 'labsubtype' should indicate 'subtype information not available'. This is corrected as 'labsubtype1' = 9999

After data cleaning, a total of 468 records were retained for further processing. Some coding errors in symptom onset data were discovered as only records were shared for which the swab was taking within one week after symptom onset (Table 3).

Table 3 Additional data quality issues, and corrections

Patient id	Data quality issue	Action
11724388:	For all these patients, it was reported that the 'onset of symptoms started a week ago'; The onset date was derived incorrectly.	Onset data corrected as:05/01/2019
11265556:	Same as above	Onset data corrected as:02/01/2019
12397808:	Same as above	Onset data corrected as:25/01/2019
11908940:	Same as above	Onset data corrected as:30/01/2019
23825:	Same as above	Onset data corrected as:19/02/2019
12246779:	Same as above	Onset data corrected as:20/01/2019
574338:	Same as above	Onset data corrected as:01/02/2019

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are applied and records with missing data in the outcome, exposure and covariate information are discarded. The following exclusion criteria are applied:

1. is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)
2. is less than 6 months of age at the time of the onset of the symptoms
3. has a contraindication for influenza vaccine
4. is institutionalised at the time of symptoms onset
5. will have the respiratory specimen taken ≥ 8 days after SARI onset
6. tested positive for any influenza virus in the current season before the onset of symptoms leading to the current primary care visit/hospitalisation
7. was previously hospitalised < 48 hours prior to SARI onset
8. had his/her ILI/SARI onset ≥ 48 hours after hospital admission

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

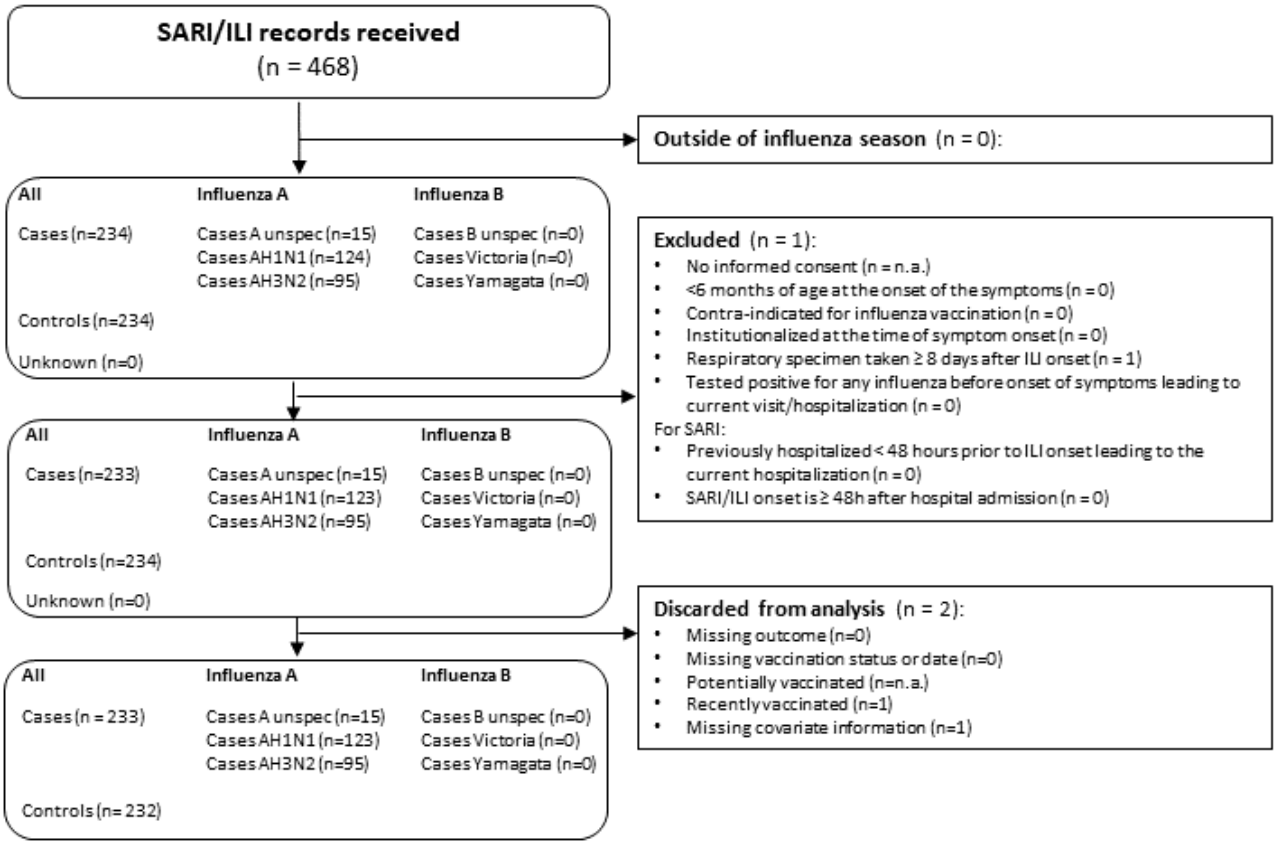


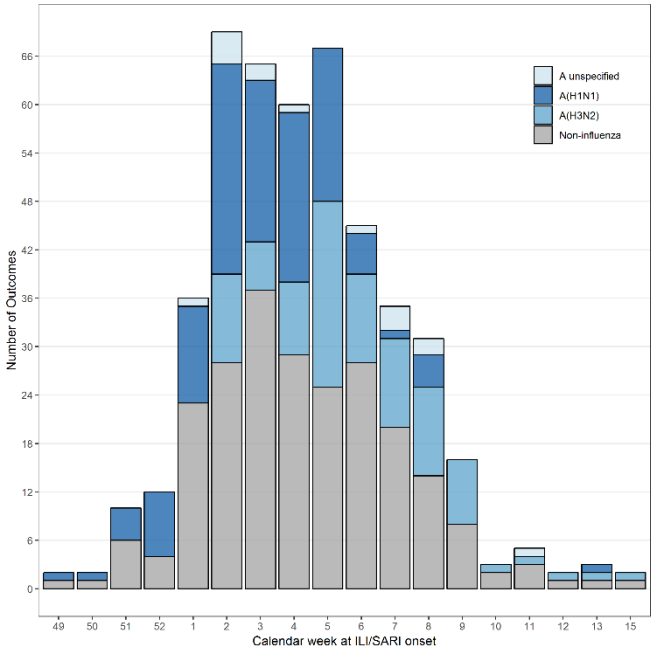
Figure 1 Attrition diagram

Data summary

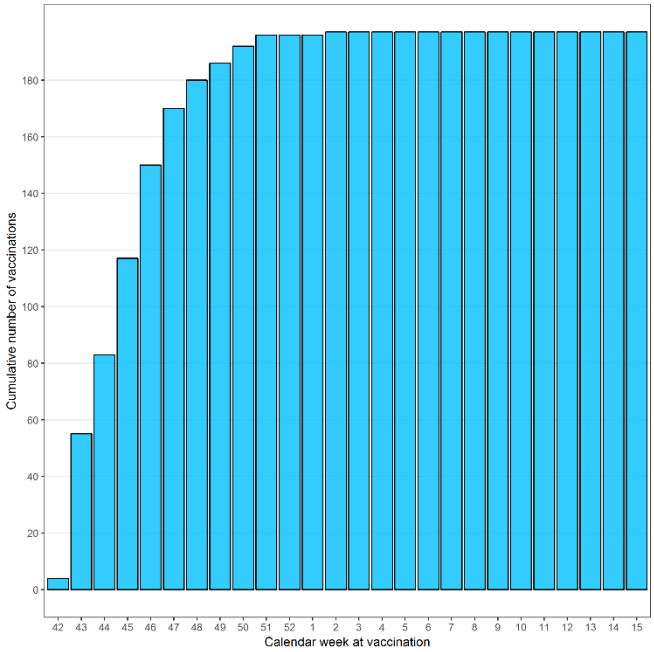
A total of 465 subjects were retained for analysis, with the majority of subjects being elderly (65+ yrs) (Table 3). Graphical summaries of the data are provided in Figure 2.

Table 3: Data for analysis: characteristics

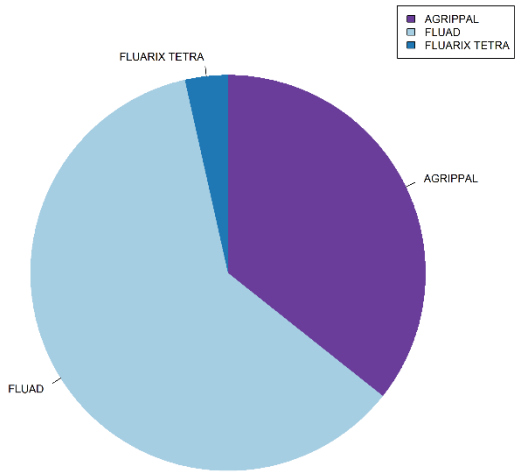
Characteristic	Number of subjects
Subjects	465
Influenza cases	233
Vaccinated subjects	197
Age	
Children (0-17y)	70
Adults (18-64y)	124
Elderly (65+y)	271
Influenza vaccine brand	
Agripal	71
Fluad	121
Fluarix Tetra	7



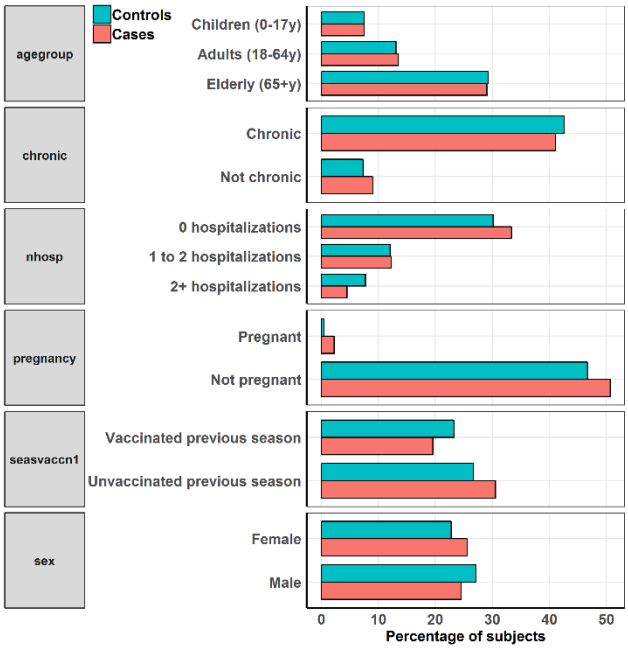
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among cases and controls

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: National Institute for Health and Welfare - THL, FINLAND

REGISTER-BASED COHORT STUDY

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Kaatje Bollaerts (P95) Jorne Biccler (P95) Ulrike Baum (THL)
Data provider	Ulrike Baum

Version	Date	Description
V0.1	6 June 2019	First Draft
V1.0	13 June 2019	Approved version

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Finnish National Institute for Health and Welfare (THL) (Italy). The report has been developed based on the following reference document:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)

THL, Finland

The Finnish National Institute for Health and Welfare (THL) participates to the DRIVE 2018/19 study with a register-based cohort study. Only aggregated data (i.e. number of events and person time within strata of interest) were shared. This implies that the data quality checks and study inclusion/exclusion cannot be carried out centrally.

The THL dataset was uploaded for ‘analysis’ using the DRIVE Electronic Study Support Application on 15/05/2019.

Data quality checks

The quality of the provided data generally corresponds to the quality of the data in the utilised, study-independently maintained registers. No particular data quality checks were performed.

Data were extracted from the respective registers and linked on the individual-level through a unique person identifier. Outcome, exposure and covariate records that could not be deterministically linked to the individuals in the study cohort (defined as all valid person identifiers in the Finnish Population Information System) were not considered in the study. Other unidentifiable records, for example data in the Finnish National Vaccination Register (NVR) without information on the administered vaccine, were also not considered. In addition, records had to be plausible by not dating an individual’s event, such as vaccination, to a time before the individual’s birth in order to be considered in the study.

Inclusion/exclusion criteria

Inclusion criteria:

- Alive on 01/10/2018
- Aged 6 months to 6 years or 65 to 100 years on 01/10/2018
- Registered in Finnish Population Information System
- Residence in Finnish municipality on 11/05/2019

Exclusion criteria:

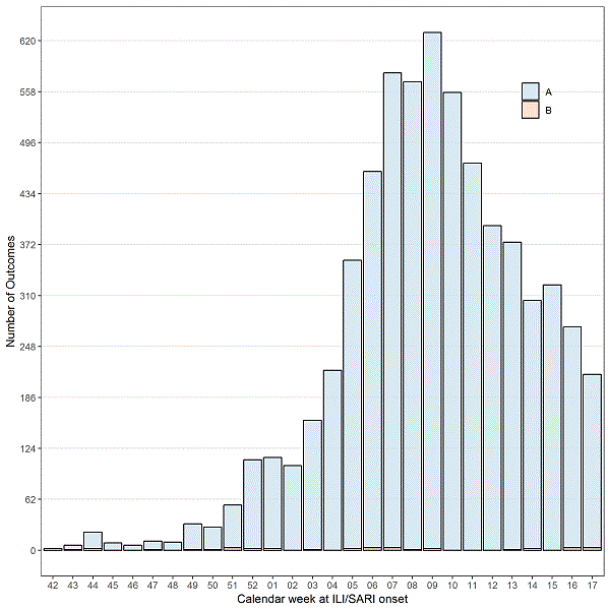
- Residence outside the NVR's catchment area between 02/10/2017 and 20/05/2018
- Residence outside the NVR's catchment area between 01/10/2018 and 11/05/2019
- Seasonal influenza vaccination between 01/08/2018 and 30/09/2018

Data summary

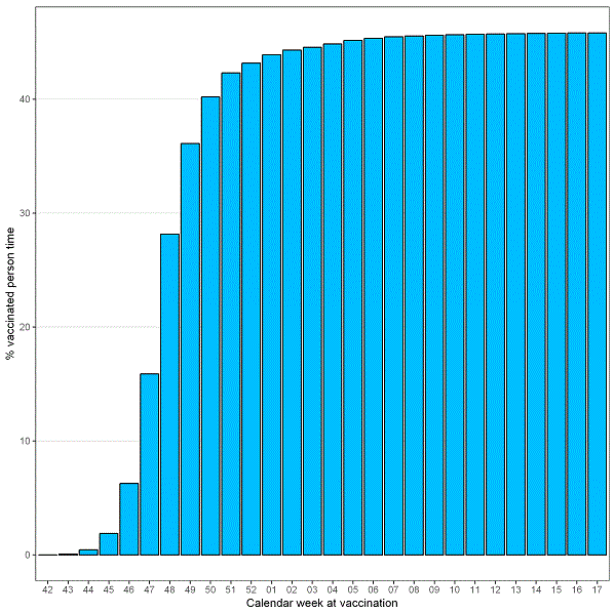
Data on the general population 6 months to 6 years (168020.7 person years) and >= 65 years (600394.9 person years) were shared. Tabular and graphical summaries of the data are provided in Table 1 and in Figure 1.

Table 1: Data for analysis: characteristics

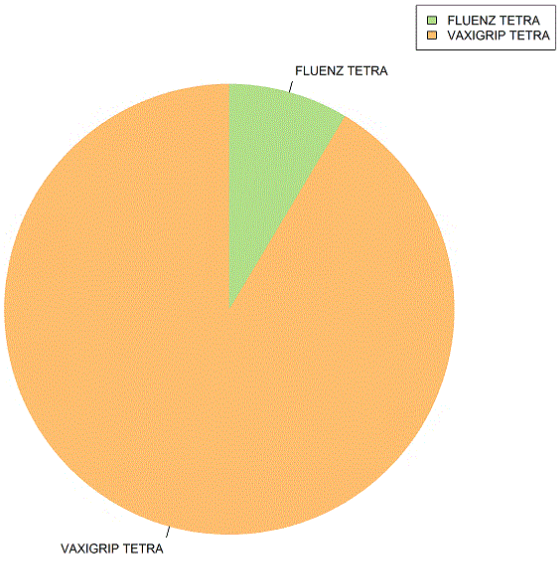
Characteristic	Number of subjects
Children (6m-6yr)	
Person years	168020.7
% vaccinated person time (Fluenz Tetra)	15.3%
% vaccinated person time (Vaxigrip Tetra)	9.8%
Influenza cases	4894
Elderly (>=65 yr)	
Person years	600394.9
% vaccinated person time (Vaxigrip Tetra)	39%
Influenza cases	9083



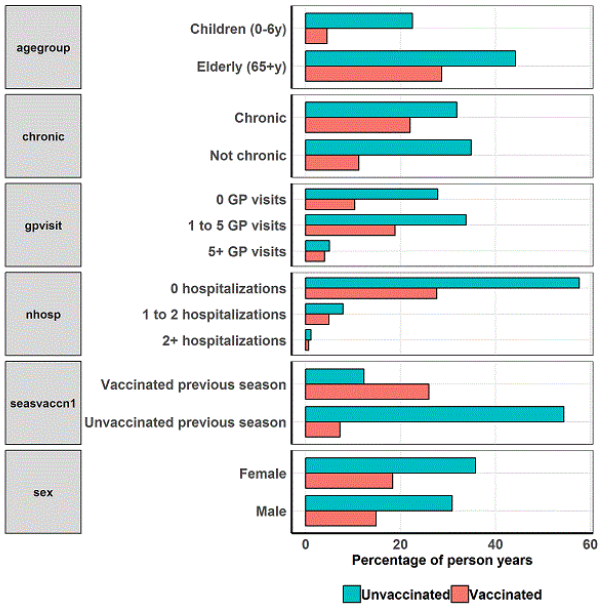
Number of influenza cases by type, over time



Percentage of vaccinated person time, by week



Distribution of vaccine brands



Distribution of covariates among exposed and unexposed

Figure 1 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili, ITALY

HEALTH CARE WORKERS COHORT STUDY

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Kaatje Bollaerts (P95) Jorne Biccler (P95)
Data provider	Stefano Mosca

Version	Date	Description
V0.1	8 June 2019	First Draft
V1.0		

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI) (Italy). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements for cohort studies

CIRI, Italy

Centro Interuniversitario di Ricerca sull'Influenza e sulle altre infezioni trasmissibili (CIRI) is a University Hospital in Genua, Italy. They participate to the DRIVE 2018/19 multi-site study with data collected using a primary care-based test-negative design (TND) study and a cohort study in health care workers (HCW). This report describes the HCW cohort data. A description of the study characteristics can be obtained from the DRIVE SAP 2018/19.

The dataset used for central data quality checks was uploaded for 'analysis' using the DRIVE Electronic Study Support Application on 15/05/2019. The uploaded TND dataset contained records on 4468 enrolled participants.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement for cohort studies are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements for cohort studies?	Yes
#3.	There are no duplicated records in the dataset?	Yes
#4	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements for cohort studies?	Yes
#5.	All records have onset date before swab date?	No, 2 records violate this condition, See Table 2
#9	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#7	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#8	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
#5	808	Swab date before onset date	Site was contacted. There was a coding error in the swab date, corrected as 'swabdate'= 2019-02-01
#5	4273	Swab date before onset date	Site was contacted. There was a coding error in the swab date, corrected as 'swabdate'= 2018-11-29

After data cleaning, a total of 4468 records were retained for further processing.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are checked. The following criteria are applied:

1. Inclusion criteria: in service prior to start of follow-up in Week 42, 2018
2. Exclusion criteria:
    - is unwilling to participate or unable to communicate and give consent
    - contra-indicated for influenza vaccination
    - have the respiratory specimen taken ≥ 8 days after ILI onset

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

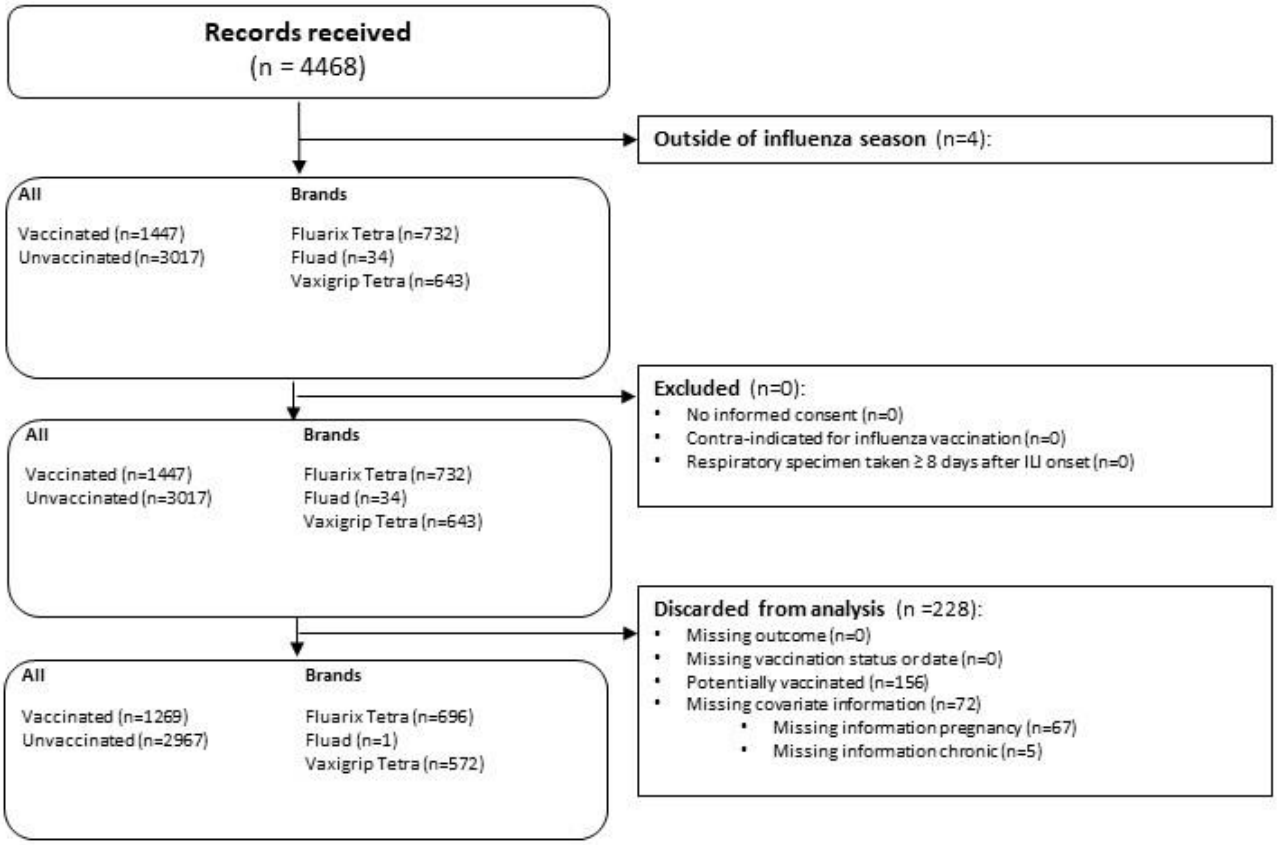


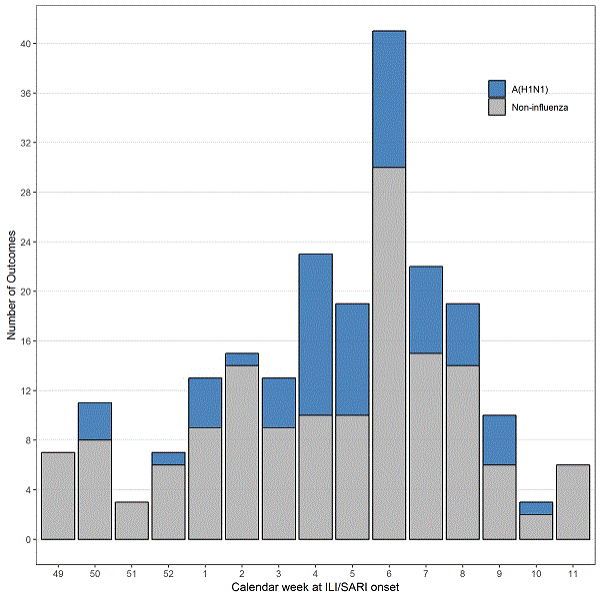
Figure 1 Attrition diagram

Data summary

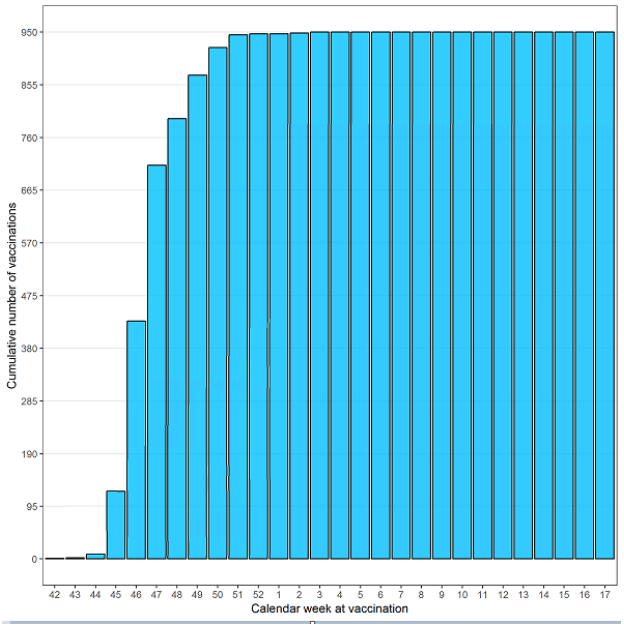
A total of 4223 subjects were retained for analysis, with Fluarix Tetra being mostly given to HCW from the Liguria region and Vaxigrip Tetra to the HCW from the Lombardia region. The average follow up time was 3.32 person months. Tabular and graphical summaries of the data are provided in Table 3 and in Figure 2.

Table 3: Data for analysis: characteristics

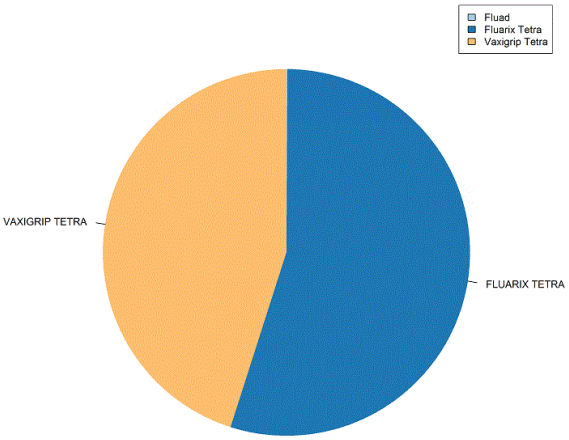
Characteristic	Number of subjects
Subjects	4236
Influenza cases	63
Vaccinated subjects	1269
Age	
Adults (18-64y)	4137
Elderly (65+y)	99
Influenza vaccine brand	
Fluarix Tetra	572
Vaxigrip Tetra	696
Fluad	1



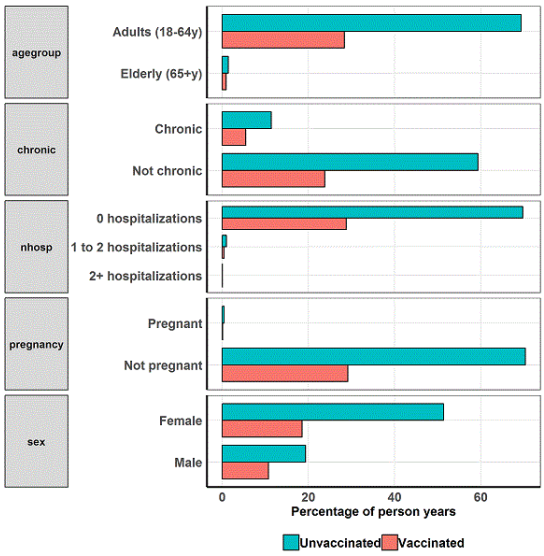
Number of non-influenza ILI cases and influenza cases by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands



Distribution of covariates among exposed and unexposed

Figure 2 Visualizations

Brand-specific influenza vaccine effectiveness in Europe, season 2018/19

DATA QUALITY REPORT: University of Athens, GREECE

PREGNANCY COHORT STUDY

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Authors	Jorne Biccler (P95) Kaatje Bollaerts (P95)
Data provider	Helena Maltezou

Version	Date	Description
V0.1	12 June 2019	First Draft
V1.0		

Background

For every site, a data quality report is produced. The report contains a description of the results of the quality checks performed, the amount of data that will be retained for analysis after applying the in-and exclusion criteria and graphical summaries of the retained data.
This report is the data quality report of the data provided for the 2018/19 influenza season by the University of Athens (UoA) (Greece). The report has been developed based on the following two reference documents:

- DRIVE Statistical Analysis Plan (SAP) season 2018/19, registered at the ENCePP EU PAS Register (EUPAS29817)
- DRIVE minimal data requirements for cohort studies

University of Athens, Greece

The University of Athens (UoA) participated to DRIVE 2018/19 with a cohort study in pregnant women and their infants. A description of the study characteristics can be obtained from the DRIVE SAP 2018/19.

The dataset used for central data quality checks was uploaded for ‘analysis’ using the DRIVE Electronic Study Support Application on 15/05/2019. The uploaded TND dataset contained records on 647 women, of which 423 women gave birth to 446 children during the study follow-up and of which 423 women were followed during the study period.

Data quality checks

Several types of quality checks were performed, including providing all mandatory variables, adherence to the variable naming convention as specified in the minimal data requirements, presence of duplicated records, variable formats and inconsistencies between variables. When serious data quality issues were found, the data site responsible person was contacted and the data were either corrected or discarded from further analysis. An overview of the data quality checks is provided in Table 1. Details on the data quality violations and subsequent actions taken are given in Table 2.

Table 1 Overview of data quality checks

Quality check Nr.	Description	Outcome
#1.	All mandatory variables as specified in the DRIVE minimal data requirement for cohort studies are included in the dataset?	Yes
#2.	All variables follow the naming conventions as specified in the DRIVE minimal data requirements for cohort studies?	Yes
#3.	There are no duplicated records in the dataset?	Yes
#4	All variables have proper variable formats and values/coding as specified in the DRIVE minimal data requirements for cohort studies?	No, the dates were not uniformly coded, see Table 2
#5.	All records have onset date before swab date?	Yes
#9	All records for which the 18/19 influenza vaccination status is positive, have a valid vaccination date?	Yes
#7	Information on influenza subtype/lineage consistent with information on influenza type (first infection)?	Yes
#8	Variable 'at least one chronic condition' consistent with information provided on individual chronic conditions (when provided)?	Yes

Table 2 Records with violations against the quality checks, details

Quality check nr	Patient id	Data quality issue	Action
4	Most patients, e.g. the startdate of the subject with id 6	Non-uniform coding of the year in the date variables	All dates were converted to the dd/mm/yyyy format

After data cleaning, a total of 869 records were retained for further processing.

Inclusion/exclusion criteria

After performing the data quality checks and implementing the corrective measures, the study /exclusion criteria are checked. The following criteria are applied:

The following in- and exclusion criteria will be applied to all study subjects;

Inclusion criteria:

- age 18 to ≤45 years
- stable health
- presented to the outpatient clinic of the department of obstetrics and gynaecology between October 1 and December 31, 2018

Exclusion criteria:

- is unwilling to participate or unable to communicate (in Greek or English) and give consent
- received influenza vaccine < 6 months prior to study entry
- received any investigational drug or product < 30 days prior to study entry
- history of Guillain-Barré syndrome
- history of hypersensitivity to influenza vaccines or its components
- immunosuppression
- received immunoglobulins or blood products < 3 months

An overview of the number of records excluded (per exclusion criterion) or discarded (per type of variable with missing information) is given in Figure 1.

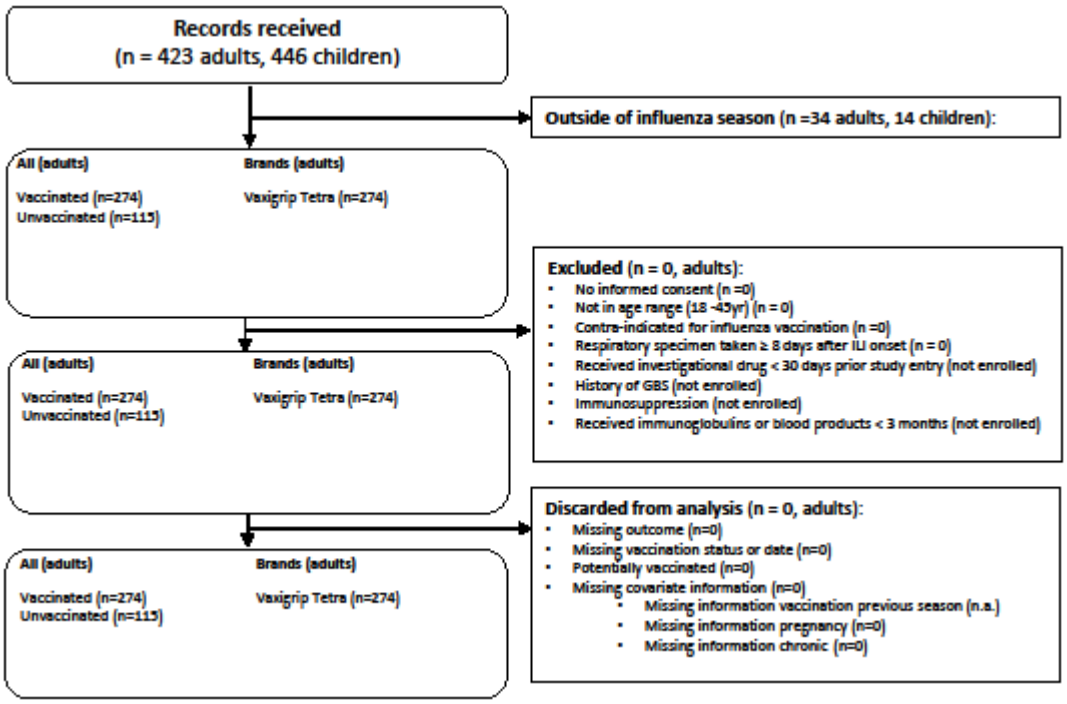


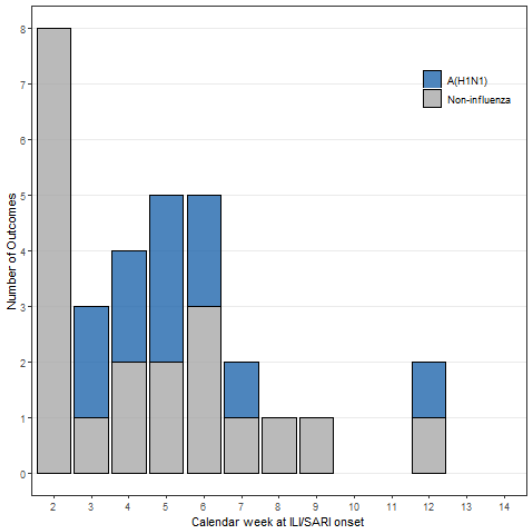
Figure 1 Attrition diagram

Data summary

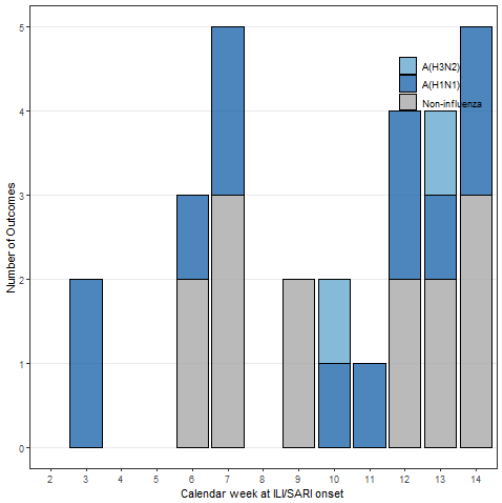
A total of 821 subjects were retained for analysis. The average follow up time of the women and infants was 2.07 and 2.30 person months. Tabular and graphical summaries of the data are provided in Table 3 and in Figure 2.

Table 3: Data for analysis: characteristics

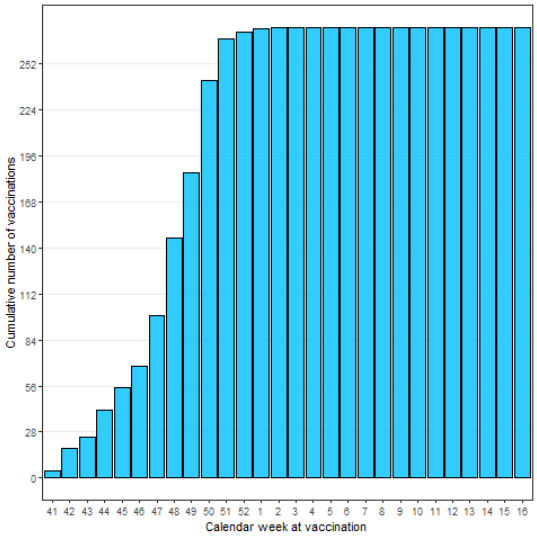
Characteristic	Number of subjects
Women	389
Influenza cases	11
Vaccinated subjects	274
Vaxigrip Tetra	274
Infants	432
Influenza cases	14
Vaccinated mothers	249



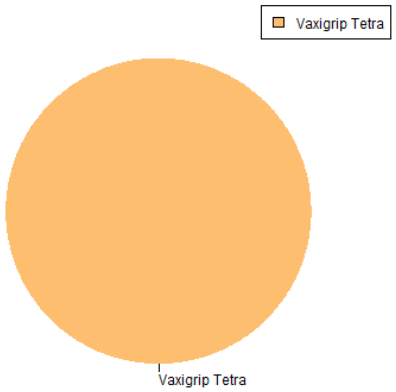
Number of non-influenza ILI cases and influenza cases among adults by type, over time



Number of non-influenza ILI cases and influenza cases among infants by type, over time



Cumulative number of vaccinations over time



Distribution of vaccine brands

Figure 2 Visualizations

Annex 4. ESSA User Manual

DRIVE ELECTRONIC STUDY SUPPORT APPLICATION (ESSA):

USER MANUAL

DRIVE
DEVELOPMENT OF ROBUST AND INNOVATIVE VACCINE EFFECTIVENESS

Application Developers	Tom De Smedt (P95) Maria Alexandridou (P95) Margarita Riera (P95) Kaatje Bollaerts (P95)
Authors User manual	Harshana Liyanage (University of Surrey) Uy Hoang (University of Surrey) Roberto Bonaiuti (University of Florence) Alfredo Vannacci (University of Florence) Maria Alexandridou (P95) Tom De Smedt (P95) Kaatje Bollaerts (P95)

Document History

Version	Date	Description
V1.0	31 Oct 2018	First Draft
V2.0	21 Dec 2018	Second Draft
V2.1	3 Jan 2019	Minor modifications to second Draft

Contents

Index of figures

Index of tables

General introduction and description of the ESSA

In the DRIVE project, influenza vaccine effectiveness (IVE) data from several independently operating national and regional study sites will be analysed together to obtain brand-specific IVE estimates.
The DRIVE Electronic Study Support Application (ESSA) facilitates the quality control of the data and the safe uploading of data to the central DRIVE research server. The DRIVE ESSA is a web application accessible from any web browser serving the following purposes:

- Aiding research sites to do the quality assurance of their data by automatically performing data quality checks
- Providing a visual summary of the data
- Allowing research sites to share the visual summaries and tables for monitoring purposes
- Allowing research sites to safely upload their data to the central DRIVE Research Server for analysis

For any questions, please contact: maria.alexandridou@p-95.com

ESSA Login process

In order to access to the DRIVE ESSA, use the following link: **http://apps.p-95.com/essa1/**.
The user can log in to the web application by using a personal email address and password combination.

Data Flow

In Figure 1 the data flow to the DRIVE Research server is detailed. The interim and final study data is uploaded by the DRIVE research study sites to the ESSA Server using a secure connection. After the data is checked, the DRIVE research study site can decide to share minimal data for monitoring or to transfer the final data to the DRIVE Research Server for analysis. All data transfers will take place using a secure connection.

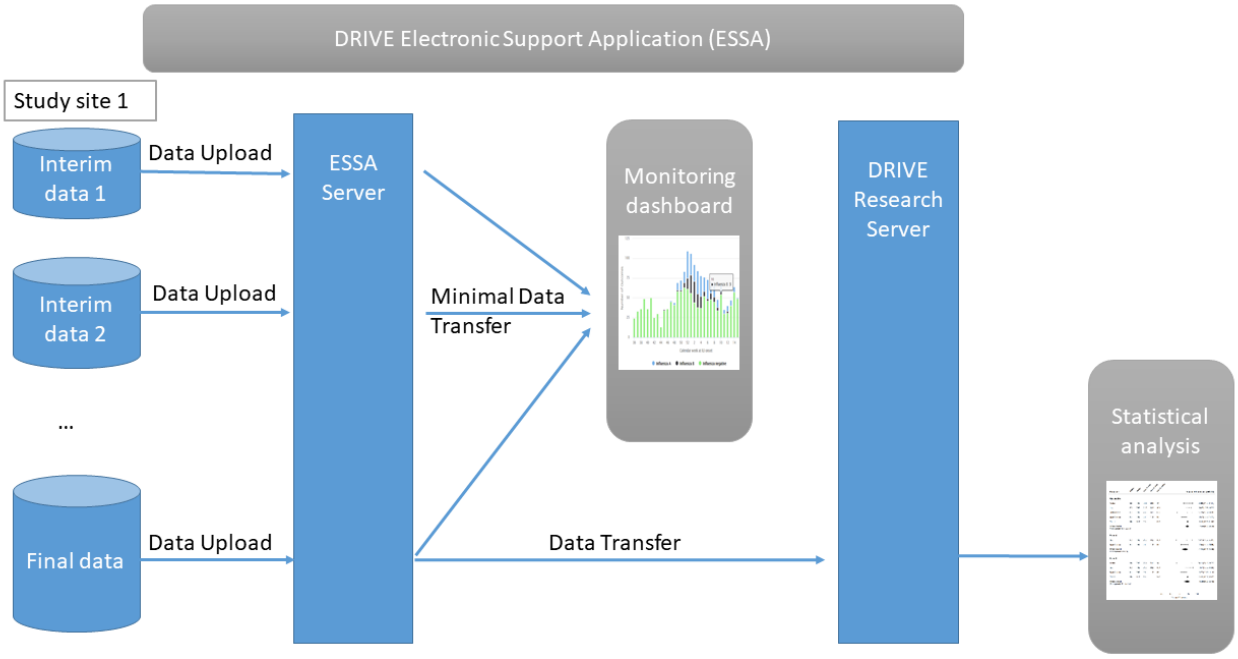


Figure 1 Data flow

Main commands of the application

The sidebar on the left of the ESSA is the main control tool of the application. The options available in the sidebar are:

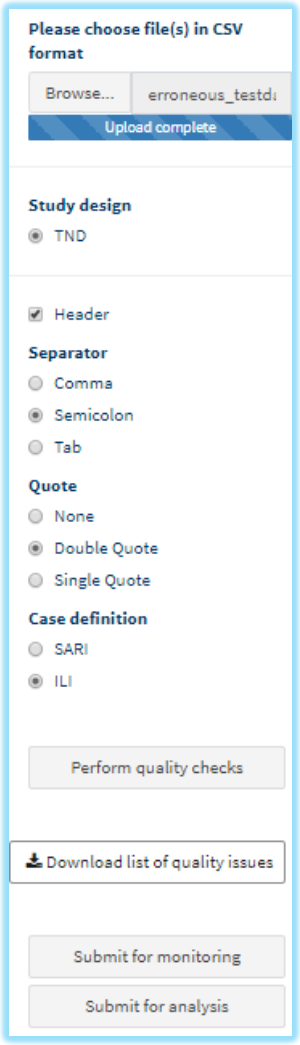


Figure 2 ESSA control side bar

- The button **Browse** to select the dataset for uploading to the ESSA;
- The **Study design** button to select the study design of the uploaded dataset. Currently, only TND (Test-Negative Design) option is available;
- The **Header** checkbox to indicate if the first row of the data contains the column headers;
- The radio buttons **Separator** to indicate whether the data fields are separated by a comma, semicolon or tab;
- The radio buttons **Quote** to indicate whether each field in the data is contained within single quotes, double quotes or no quotes are used;
- The radio buttons **Case definition** to indicate whether the SARI (Severe Acute Respiratory Infection) or ILI (Influenza Like Illness) case definition was used for the uploaded dataset;
- The action button **Perform quality checks** to perform the quality checks of the data;
- The action button **Download list of quality issues** to download a list with quality issues found through performing the quality checks.
- The action button **Submit for monitoring** to allow sharing the data summaries (tables and figures) with the DRIVE study team;
- The action button **Submit for analysis** to indicate the uploaded data can be used for analysis;

Step 1: upload a dataset to the ESSA

Dataset format

Datasets must be in CSV format (with no size limit) and be compliant with the minimum data requirements (link). Some variables are obligatory, other optional. The compliance to the minimum data requirements will be checked by the ESSA application.

Upload data procedure

1. Click the **Browse** button located at the top of the Control Sidebar of the ESSA (

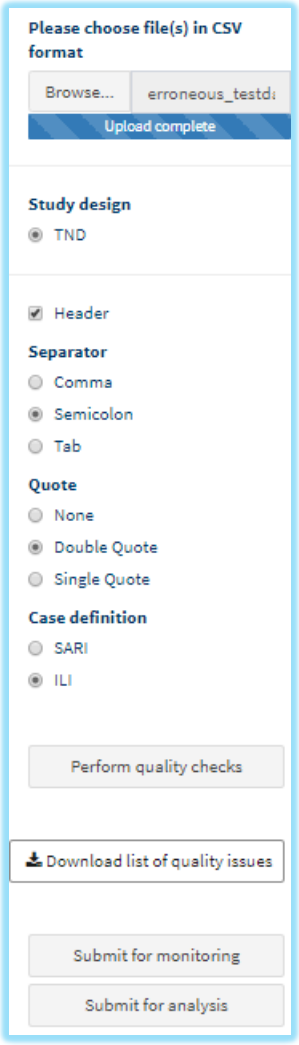


Figure 2 ESSA control side bar**Fout! Verwijzingsbron niet gevonden.**).

2. Browse computer folders and select the .csv file to upload. The file will be uploaded to the ESSA.

3. Indicate how the data should be read using the **Study design**, **Header**, **Separator**, **Quote** and **Case definition** buttons.

4. Preview the uploaded data and change the options above if the data are not properly read. See examples below.

An example of a properly uploaded dataset is given in Figure 3 . An example of data that is not properly uploaded is given in Figure 4. In this example view, the following problems are present and can be solved using the options available in the ESSA control sidebar:

Problem	How to solve
Variables names are in the first data row of the dataset	Check the **Header** checkbox in the sidebar
Variables are not placed in columns	Select among the **Separator** options, the proper field separator (“Comma” in this case)
Variables values appear in quotes	Select among the **Quote** options, how the fields are contained (“Double Quote” in this case)

Show 10 entries

X	idcountry	id	region	gp	hosp	sex	dob	onsetdate
1	SP	1000001	Valencia		10	1	15/11/1949	07/09/2017
2	SP	1000002	Valencia		10	0	01/04/1923	08/09/2017
5	SP	200003	Valencia		2	1	18/09/1942	09/09/2017
6	SP	200005	Valencia		2	0	30/06/1971	05/09/2017
12	SP	700005	Valencia		7	1	23/01/1975	09/09/2017
14	SP	1000007	Valencia		10	1	08/12/1940	07/09/2017
16	SP	1000009	Valencia		10	0	11/04/1934	10/09/2017
22	SP	200010	Valencia		2	0	04/02/1934	07/09/2017

Figure 3 View of data that is properly uploaded

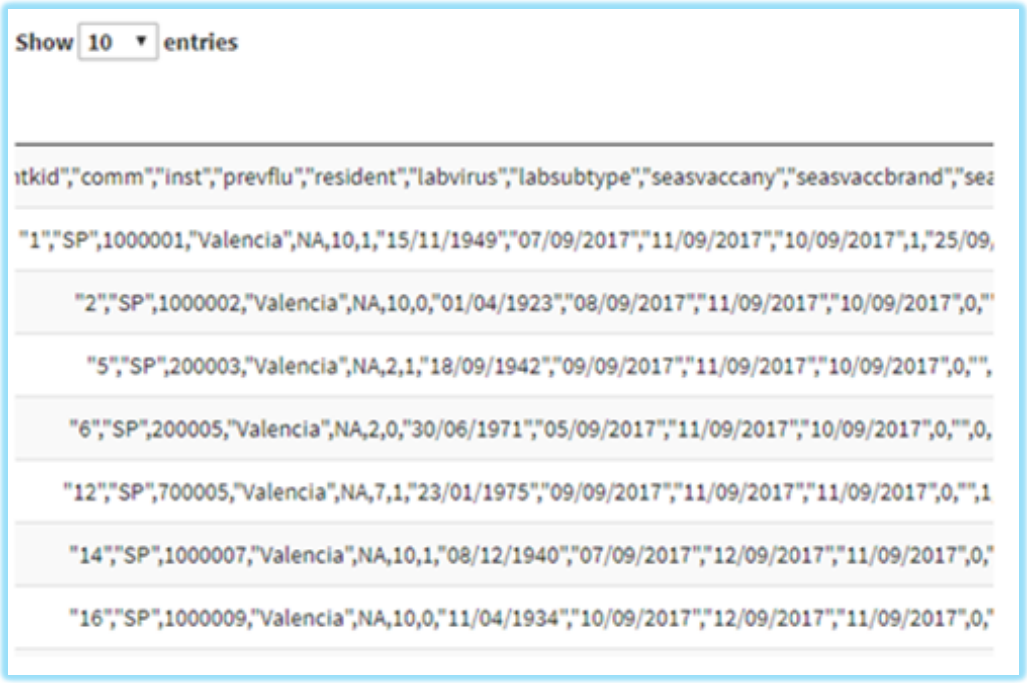


Figure 4 View of data that is not properly uploaded

Step 2: Perform the data quality checks

Data quality checks are performed upon pressing the action button **Perform quality checks** (

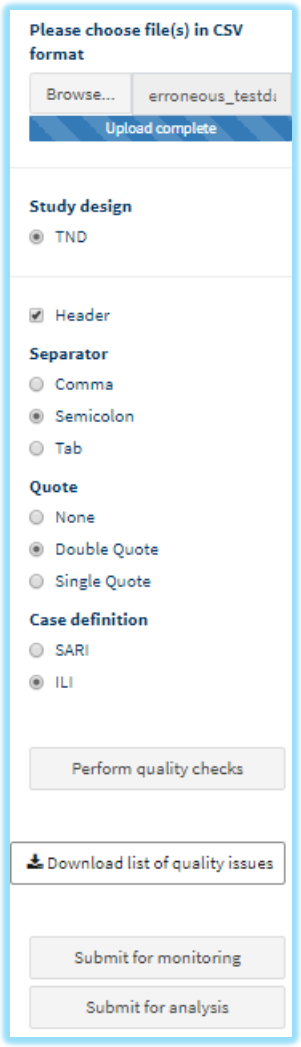


Figure 2).

The results of the quality checks will be summarized on the screen. See Figure 5 and Figure 6 for the data quality summaries of data without and with data quality issues.

Checks 1: Dataset characteristics and compliance with data requirements

- Dataset read properly.
- Dataset contains 49 rows and 74 columns.
- Dataset contains all compulsory variables.
- Dataset follows naming conventions.

Checks 2: Duplicates

- Dataset contains unique patient identifiers.
- Dataset does not contain duplicated records (i.e. different patient IDs containing same information).

Checks 3: Variables Format

These checks are restricted to ILI/SARI cases.

- No variable format issues found
- All date variables have proper date format (dd/mm/yyyy)

Checks 4: Inconsistencies between variables

These checks are restricted to ILI/SARI cases.

- All ILI cases were confirmed based on clinical symptoms (if available): TRUE
- Swab date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE
- Swab date occurs on or after visit day for all ILI/SARI cases: TRUE
- Visit date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE
- ILI/SARI onset date occurs after vaccination date for all ILI/SARI cases: TRUE
- All vaccinated ILI/SARI cases have a vaccination date and vice versa: TRUE
- Virus type and sub-type are consistent for all influenza cases: TRUE
- All ILI/SARI cases with >=1 chronic condition were confirmed based on the type of chronic condition (if available): TRUE

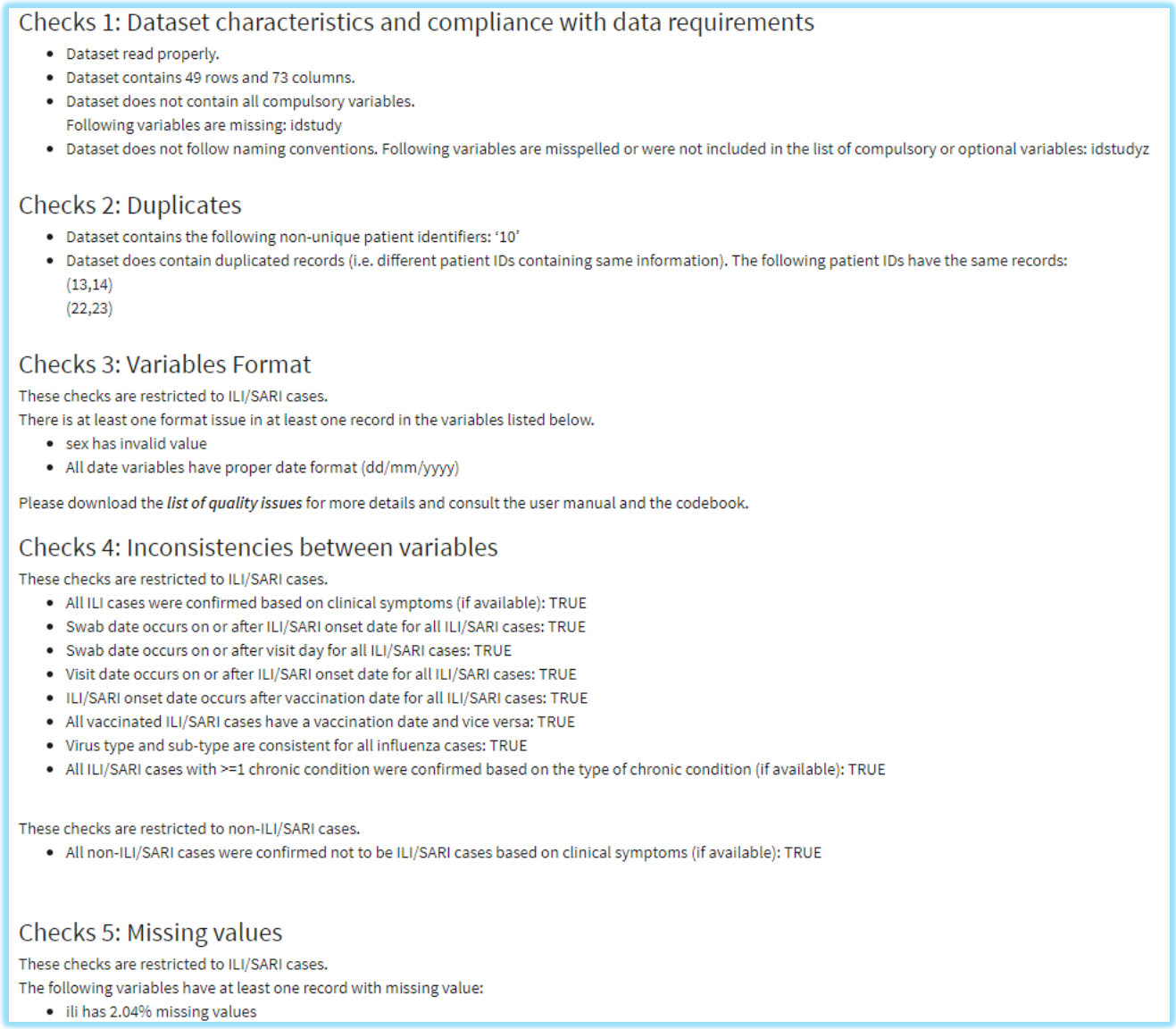
These checks are restricted to non-ILI/SARI cases.

- All non-ILI/SARI cases were confirmed not to be ILI/SARI cases based on clinical symptoms (if available): TRUE

Checks 5: Missing values

These checks are restricted to ILI/SARI cases.
All variables are complete

Figure 5 Data quality summary: data without quality issues



Checks 1: Dataset characteristics and compliance with data requirements

- Dataset read properly.
- Dataset contains 49 rows and 73 columns.
- Dataset does not contain all compulsory variables.
  Following variables are missing: idstudy
- Dataset does not follow naming conventions. Following variables are misspelled or were not included in the list of compulsory or optional variables: idstudyz

Checks 2: Duplicates

- Dataset contains the following non-unique patient identifiers: '10'
- Dataset does contain duplicated records (i.e. different patient IDs containing same information). The following patient IDs have the same records:
  (13,14)
  (22,23)

Checks 3: Variables Format

These checks are restricted to ILI/SARI cases.
There is at least one format issue in at least one record in the variables listed below.

- sex has invalid value
- All date variables have proper date format (dd/mm/yyyy)

Please download the ***list of quality issues*** for more details and consult the user manual and the codebook.

Checks 4: Inconsistencies between variables

These checks are restricted to ILI/SARI cases.

- All ILI cases were confirmed based on clinical symptoms (if available): TRUE
- Swab date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE
- Swab date occurs on or after visit day for all ILI/SARI cases: TRUE
- Visit date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE
- ILI/SARI onset date occurs after vaccination date for all ILI/SARI cases: TRUE
- All vaccinated ILI/SARI cases have a vaccination date and vice versa: TRUE
- Virus type and sub-type are consistent for all influenza cases: TRUE
- All ILI/SARI cases with >=1 chronic condition were confirmed based on the type of chronic condition (if available): TRUE

These checks are restricted to non-ILI/SARI cases.

- All non-ILI/SARI cases were confirmed not to be ILI/SARI cases based on clinical symptoms (if available): TRUE

Checks 5: Missing values

These checks are restricted to ILI/SARI cases.
The following variables have at least one record with missing value:

- ili has 2.04% missing values

Figure 6 Data quality summary: data with quality issues

Checks 1: Dataset characteristics and compliance with data requirements

The ESSA will provide an indication about the status of the upload and inform the user about the number of rows and columns processed in the uploaded data file.

The data will be validated with respect to the minimal data requirements specified by the DRIVE codebook. If there are compulsory variables missing in the variable names, the ESSA will provide an error message ("Dataset does not contain all compulsory variables") and provide a list of missing variables in the data uploaded. The ESSA will also check if the variable names are consistent with respect to the naming conventions specified in the DRIVE codebook. It will indicate any misspelt or redundant variable names.

Checks 2: Duplicates

The ESSA will check whether there are duplicated patient IDs (patient identifiers) and whether different patient IDs contain the same information (i.e. all the information is the same apart from the patient IDs). If there are any duplicated patient IDs, the DRIVE ESSA will show the duplicated patient IDs. If there are multiple patient IDs that contain the same information, the DRIVE ESSA will show these multiple patient IDs sharing the same information between round brackets.

Checks 3: Variables Format

The variables are checked with respect to the acceptable format and value ranges as specified in the DRIVE codebook. For example, valid values for the variable “fever” are 0 (no), 1 (yes) and 9999 (no information). Other values will generate a data quality issue and will be mentioned in the data quality summary. Details on the variable format issues will be given in the CSV file that can be downloaded using the action button in the sidebar **Download list of quality issues.**

Checks 4: Inconsistencies between variables

Several consistency checks between variables are performed. The different consistency checks are described in Table 1: Inconsistency checks and Table 2. If a check cannot be done because variables are not included in the data set, then the message ‘Check not possible’ will be given. Details on the inconsistencies found will be given in the CSV file that can be downloaded using the action button in the sidebar **Download list of quality issues.**

Table 1: Inconsistency checks for the ILI/SARI cases only

All ILI cases were confirmed based on clinical symptoms (if available): TRUE/FALSE

(activates only if the ILI case definition is chosen on the side bar)

Checks if every ILI patient does follow the case definition (Sudden onset & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are ILI patients whose symptoms do not indicate that they have ILI, then this is displayed on the data quality summary as FALSE. If the ILI symptoms are not included in the submitted data set, then the default message is TRUE.

All SARI cases were confirmed based on clinical symptoms (if available): TRUE/FALSE

(activates only if the SARI case definition is chosen on the side bar)

Checks if every SARI patient does follow the case definition (Deterioration & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are SARI patients whose symptoms do not indicate that they have SARI, then this is displayed on the data quality summary as FALSE. If the SARI symptoms are not included in the submitted data set, then the default message is TRUE.

Swab date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of swab is on the same date or after the date of onset.

Swab date occurs on or after visit day for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of swab is on the same date or after the date of visit.

Visit date occurs on or after ILI/SARI onset date for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of visit is on the same date or after the date of ILI/SARI onset.

ILI/SARI onset date occurs after vaccination date for all ILI/SARI cases: TRUE/FALSE

Checks whether the date of ILI/SARI onset is on the same date or after the date of vaccination.

All vaccinated ILI/SARI cases have a vaccination date and vice versa: TRUE/FALSE

Checks whether the vaccinated persons have a vaccination date and vice versa.

Virus type and sub-type are consistent for all influenza cases: TRUE/FALSE

Checks if virus type and virus subtype are consistent. (e.g. a patient with influenza A and subtype B Yamagata would be flagged as inconsistent)

All ILI/SARI cases with >=1 chronic condition were confirmed based on the type of chronic condition (if available) : TRUE/FALSE

Checks if the chronic case definition can be confirmed using at least one of the following conditions: chronic liver disease, diabetes, cardiovascular diseases, cancer, immune-deficiencies, lung disease, anaemia, renal disease, dementia, stroke, rheumatoid disease, obesity, perinatal/congenital risk factor. If chronic symptoms are not included in the submitted data set, then the default message is TRUE.

Table 2: Inconsistency checks for patients that are non-ILI/SARI cases

All non-ILI cases were confirmed not to be ILI cases based on clinical symptoms (if available): TRUE/FALSE Checks if every non-ILI patient does not follow the ILI case definition (Sudden onset) & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are non-ILI patients whose symptoms indicate that they have ILI, then this is displayed on the screen as false. If the ILI symptoms are not included in the submitted data set, then the default message is true.
All non-SARI cases were confirmed not to be SARI cases based on clinical symptoms (if available): TRUE/FALSE Checks if every non-SARI patient does not follow the SARI case definition (Deterioration) & (Fever OR Headache OR Myalgia OR Malaise) & (Cough OR Difficulty breathing OR Sore throat)). If there are non-SARI patients whose symptoms indicate that they have SARI, then this is displayed on the screen as false. If the SARI symptoms are not included in the submitted data set, then the default message is true.

Checks 5: Missing values

The ESSA checks for each variable the amount of missing data. For each variable with missing data, the percentage of missing data will be given in the data quality summary.

Step 3: Download the list with quality issues found

Upon performing the quality checks, a list with identified quality issues will be automatically generated. A CSV file with the list of identified quality issues can be downloaded by pressing the action button in the sidebar **Download list of quality issues**. Figure 7 gives an excerpt of such a CSV file with identified quality issues. A record is generated for every patient ID for which at least one quality issue was found.

ID	issues
1	Row issue(s): - swabdate occurs earlier than onsetdate
2	Row issue(s): - agemonths has invalid value
3	Row issue(s): - agemonths has invalid value
7	Row issue(s): - ILI case definition cannot be verified based on symptoms
8	Row issue(s): - bmi has negative value
12	Row issue(s): - virus type not consistent with subtype
14	Row issue(s): - idcountry is not unique
17	Row issue(s): - onsetdate occurs earlier than vaccination date - virus type not consistent with subtype
20	Row issue(s): - death has invalid value
21	Row issue(s): - smoking has invalid value
23	Row issue(s): - vaccination status not consistent with vaccination date
28	Row issue(s): - chronic case definition inconsistent with symptoms
32	Row issue(s): - vaccination status not consistent with vaccination date
33	Row issue(s): - consentkin has invalid value
49	Row issue(s): - setting has invalid value - seasvaccn2 has invalid value

Figure 7 CSV file with quality issues

Step 4: Explore data summaries

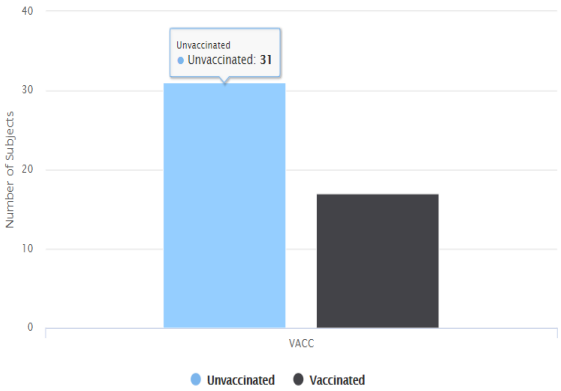
If the required information is provided, the ESSA generates three tables; one giving information on the number of subjects; one giving information on the number of brands and one giving information on the number of influenza cases by age group (Figure 8). Graphical summaries of the data are provided as well, regarding exposure (Figure 9), the outcome variables (Figure 10) and covariates (Figure 11).

	N
Nr. Subjects	49
Nr. Influenza cases	32
Nr. Vaccinated subjects	17

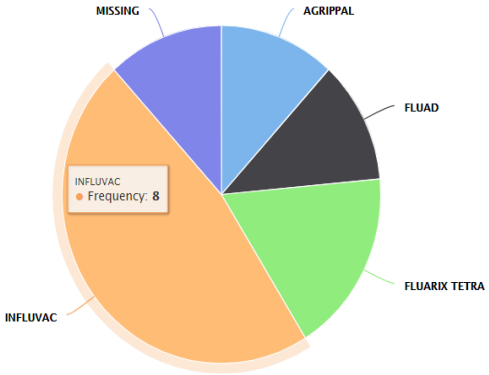
Brand	Nr. Subjects
AGRIPPAL	2
FLUAD	4
FLUARIX TETRA	3
INFLUVAC	8

Age	Nr. Influenza Cases
Children (0-14y)	10
Adults (15-64y)	14
Elderly (65+y)	8

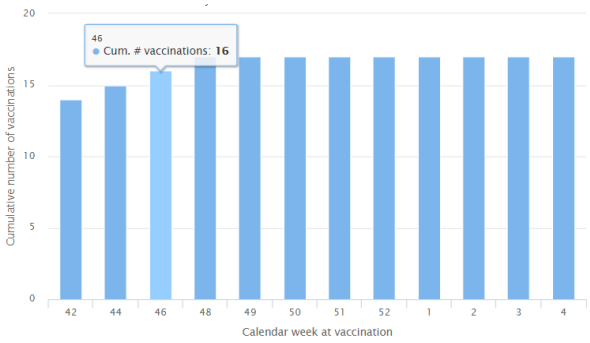
Figure 8 Data summary tables provided by ESSA



Number of (un)vaccinated subjects

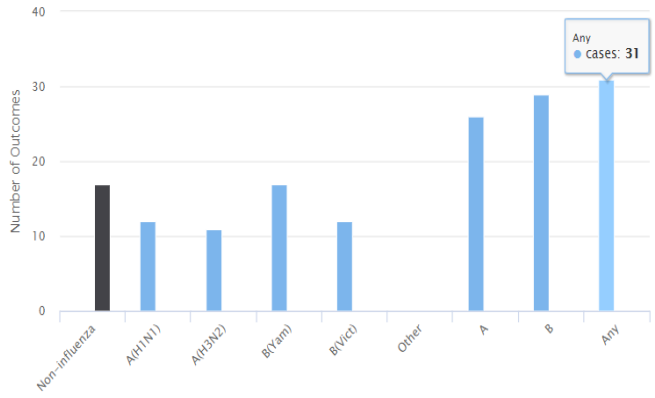


Distribution of vaccine brands

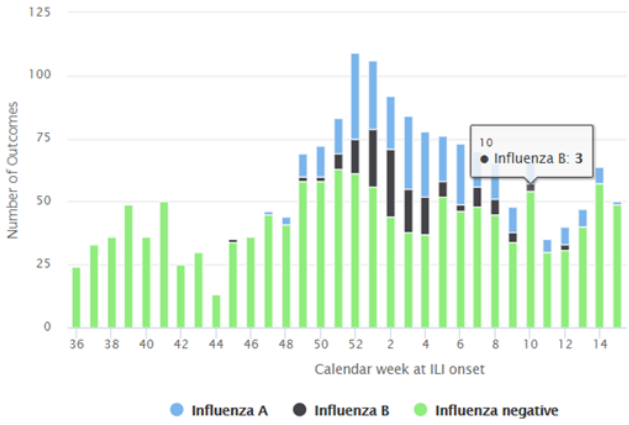


Cumulative number of vaccinated subjects over time

Figure 9 Visualizations regarding exposure

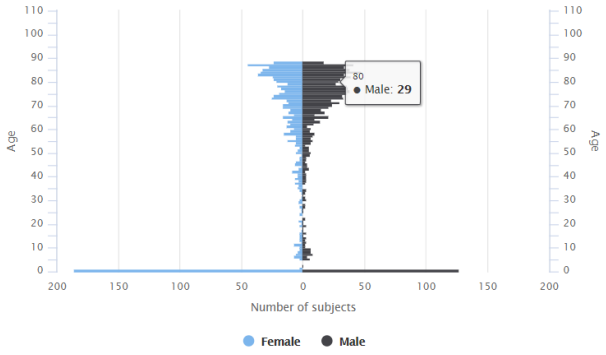


Number of cases and controls with laboratory confirmed influenza by type and subtype

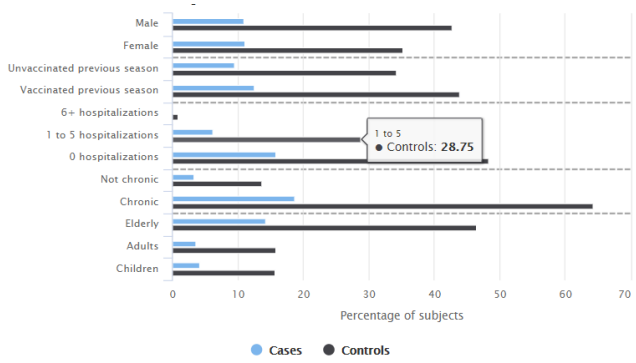


Number of laboratory-confirmed influenza infections by type over time

Figure 10 Visualizations regarding the outcome variables



Age and gender pyramid



Distribution of covariates among cases and controls

Figure 11 Visualization regarding covariates

Step 5: Share data for monitoring

After performing the quality checks and possibly modifying the data, the data summary (Tables and Figures as described in Step 4) can be shared with the DRIVE study team for data monitoring purposes. For sharing the tables and figures, press the action button **Submit for monitoring.** Sharing of data summaries for monitoring purposes can be repeatedly done throughout the influenza season.

The summary tables and figures shared by all study sites are accessible from **http://apps.p-95.com/essa2/**. The information is only accessible to members of the DRIVE study team.

Step 6: Submit data for analysis

To submit the final and clean data for analysis, press the action button **Submit for analysis.**

Annex 5. Dataset Metadata

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	Register based cohort study to measure brand-specific influenza vaccine effectiveness during season 2018-2019	Name of the dataset or research project that produced it
Data owner:	Finnish Institute for Health and Welfare (THL), Mannerheimintie 166, 00300 Helsinki, Finland	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 01.10.2018 End Date: 19.05.2019	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 01.10.2018 End date: 28.04.2019	Start and end date of the period covered by the dataset
Methods:	Register data deterministically linked using unique person identifier: Population Information System, National Vaccination Register, National Infectious Diseases Register, Care Register for Health Care, Register of primary health care visits. Cohort study. Software: R 3.5.3.	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Aggregated data containing only non-sensitive information	Datasets containing personal data Datasets containing non-personal data
Processing:	Aggregation by age, sex, influenza vaccination in 2017/18, presence of at least one chronic condition, number of hospitalisations and number of primary care consultations, calendar week and exposure status	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	National infectious diseases register, national vaccination register, Finnish population register	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	DRIVE and THL	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:	As described above. Number of events and person-time aggregated by the outlined covariates.	Keywords or phrases describing the subjects or content of the data
Language:	English	All languages used in the dataset
Database size	84MB	Indication of the size of the complete database

Technical description		
File inventory:	pilot2_aggregation.csv	All files associated with the project, including extensions
File formats:	csv	Format of the file
File structure:	age • sex • seasvacn1 • chronic1 • nhosp • nprim • week • exposure • events • inpatientevents • persondays • population • exp • out	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	Register holders	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Per request to register holder	Where and how your data can be accessed by other researchers
Data sharing	Per request to register holder	Description of how data will be shared, including access procedures
Ethics and legal issues	Study has been approved by register holders	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	Royal College of General Practitioners Research and Surveillance Centre (RCGP RSC) sentinel network database	Name of the dataset or research project that produced it
Data owner:	Royal College of General Practitioners Research and Surveillance Centre, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom Data custodian: Prof Simon de Lusignan	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 2004 End Date: Ongoing data collection	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 01/03/2019 End date: 28/04/2019	Start and end date of the period covered by the dataset
Methods:	Dataset is a merged output of data received from two data sources: 1) Primary data collection - Data from Point of care testing (POCT) machines: GP study sites used POCT machines to test nasal swabs for influenza. The results stored in the POCT machines were exported and sent weekly to the RCGP RSC through secure email. 2) Secondary data collection - Data from weekly updated data feed from GP study sites: RCGP RSC receives weekly data increments from practices participating in the RCGP RSC sentinel network. The data contains all clinical events added to the GP system during the previous week. Data management: MS SQL Server Ver 14	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)

Datasets containing personal data	Datasets containing personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Pseudonymised patient identifier was used in the dataset.	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	RCGP RSC sentinel network database	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	DRIVE	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:	Primary care data, England	Keywords or phrases describing the subjects or content of the data
Language:	English	All languages used in the dataset
Database size	Database currently has >4 million patients	Indication of the size of the complete database
Technical description		
File inventory:	Dataset submitted through DRIVE ESSA	All files associated with the project, including extensions
File formats:	Dataset submitted as a CSV file	Format of the file
File structure:	Dataset was prepared according to the structure specified in the DRIVE ESSA user manual	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	Data use by authorisation of RCGP RSC only	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Requests will be considered on an individual basis. See https://www.rcgp.org.uk/rsc	Where and how your data can be accessed by other researchers
Data sharing	Aggregated data can be shared once authorised by the RCGP RSC information governance committee	Description of how data will be shared, including access procedures
Ethics and legal issues	Ethical approval was obtained for the study which produced the dataset	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	Jorvi DRIVE TND 2018-19. The DRIVE dataset from the study *'The effectiveness of seasonal influenza vaccine against laboratory-confirmed influenza infection requiring hospitalization among adults in Jorvi Hospital during influenza season 2018-2019 - a test-negative case-control study'*	Name of the dataset or research project that produced it
Data owner:	HOSPITAL DISTRICT OF HELSINKI AND UUSIMAA (HUS) administrative offices at Stenbäckinkatu 9, 00029 Helsinki,Finland Institute of Health and Welfare (THL) Address: Mannerheimintie 166, PL 30, 00271 Helsinki	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 23.11.19 End Date: 30.4.19	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 23.11.19 End date: 30.4.19	Start and end date of the period covered by the dataset
Methods:	Primary data collection in prospective TND hospital setting. Secondary use of data from national health registers such as national vaccination register.	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Dataset contains personal data.	Datasets containing personal data Datasets containing non-personal data
Processing:	Data has been pseudonymised, partly checked for discrepancies and detected errors corrected, if possible	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	Source data collected on paper case report forms (CRFs)held by HUS, collected by interviews of the patient/next of kin and from hospital and municipal health centre medical records, Kanta archive and national vaccination register (NVR).	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	DRIVE-collaboration, HUS and the Finnish Institute of Health and Welfare (THL)	Organizations or agencies who funded the research, or indicate that the data owner funds the study

Content description		
Subject:	Adult, hospitalised, SARI, influenza	Keywords or phrases describing the subjects or content of the data
Language:	English	All languages used in the dataset
Database size	293	Indication of the size of the complete database
Technical description		
File inventory:	Files with study data according to the DRIVE criteria of SARI according to the DRIVE core TND protocol. (Administrative files as required by the HUS and THL internal regulations)	All files associated with the project, including extensions
File formats:	CSV format. (Administrative data: as required by HUS and THL internal regulations.)	Format of the file
File structure:	According to the DRIVE TND codebook, in compliance with DRIVE variable names.	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	Intellectual properties belong to HUS and THL as agreed in their mutual agreement. This agreement prohibits delivery of the data for research purposes to third parties other than to the DRIVE collaboration. Data use is also restricted by the protocol and the informed consent of the study subjects.	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	In aggregated form by application, if both HUS and THL accept the application.	Where and how your data can be accessed by other researchers
Data sharing	For DRIVE collaboration, the data is shared through the DRIVE ESSA application; for other partners (see above), as agreed.	Description of how data will be shared, including access procedures
Ethics and legal issues	The protocol has received a favourable opinion from the local Ethics committee (HUS district) and institutional study permission from HUS, THL and municipalities of Espoo, Kirkkonummi and Kauniainen.	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

Interuniversity Research Center on Influenza and other Transmissible Infections (CIRI-IT)

General Overview		Explanation
Title:	Case control study (test-negative design studies) to measure type/brand-specific seasonal influenza vaccine effectiveness against laboratory confirmed influenza cases in Italy, season 2018/19	Name of the dataset or research project that produced it
Data owner:	Interuniversity Research Center on Influenza and other Transmissible Infections (CIRI-IT) - University of Genoa, Via Pastore, 1 16132 Genova.	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 15-10-2018 End Date: 29-04-2018	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 15-10-2018 End date: 15-10-2018	Start and end date of the period covered by the dataset
Methods:	Study design: TND primary care Primary care data collection: data obtained by general practitioners and paediatricians, vaccination registry. Data collection: electronic and paper RCF Data management: filemaker server DB.	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Datasets containing non-personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Patients are coded as: CCSYRRRNNNN, 2 digits for City, 1 digit for Study. 1 digit for Year, 3 digit for Random letters. 4 digit for enrolment Number. Codes are pre generated for every enrolment kit.	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	Primary care (practitioners and paediatricians) data, Vaccination registry	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	FISABIO	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:		Keywords or phrases describing the subjects or content of the data
Language:	Italian, English	All languages used in the dataset
Database size	256Kb	Indication of the size of the complete database
Technical description		
File inventory:	CIRI-IT_TstNeg_14052019.csv	All files associated with the project, including extensions
File formats:	csv	Format of the file
File structure:	According to DRIVE TND case-control dataset requirements	Organization of the data file(s) and layout of the variables, where applicable
Access		

Rights	RGDP	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Data accessed CIRI-IT staff only	Where and how your data can be accessed by other researchers
Data sharing	No sharing. On demand with DRIVE	Description of how data will be shared, including access procedures
Ethics and legal issues	No problem	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	Influenza sentinel surveillance, Luxembourg, season 2018/19	Name of the dataset or research project that produced it
Data owner:	Laboratoire National de Santé Department of Microbiology 1, rue Louis Rech L-3555 Dudelange Luxembourg	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 1st Octover 2018 End Date: 19th May 2019	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 1st October 2018 End date: 19th may 2019	Start and end date of the period covered by the dataset
Methods:	Data collected by primary care physicians (GPs and paediatricians), test negative design, DNA extraction using EasyMag, testing using Seegene Allplex Respiratory Panel I on BIORAD CFX96	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Pseudonymized personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Patient and GP identifiers Identifiers were pseudonymized and are only available to PI	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:		Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:		Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:	Date, age, vaccination status, influenza test status, onset date	Keywords or phrases describing the subjects or content of the data
Language:	English	All languages used in the dataset
Database size	111kb	Indication of the size of the complete database
Technical description		

File inventory:	Drive_lux_lns1819	All files associated with the project, including extensions
File formats:	Csv file	Format of the file
File structure:	Comma separated file	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	-	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Open access	Where and how your data can be accessed by other researchers
Data sharing	Data shared once a year with study coordinator via secure web platform	Description of how data will be shared, including access procedures
Ethics and legal issues	None	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

Interuniversity Research Center on Influenza and other Transmissible Infections (CIRI-IT)

General Overview		Explanation
Title:	Population-based database cohort study to measure type/brand-specific seasonal influenza vaccine effectiveness against laboratory confirmed influenza cases in Italy, season 2018/19	Name of the dataset or research project that produced it
Data owner:	Interuniversity Research Center on Influenza and other Transmissible Infections (CIRI-IT) - University of Genoa, Via Pastore, 1 16132 Genova.	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 9-10-2018 End Date: 29-04-2018	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 9-10-2018 End date: 15-10-2018	Start and end date of the period covered by the dataset
Methods:	Study design: Clinical prospective cohort study Data collection: data obtained by face to face interview and by vaccination registry. Data collection: electronic and paper RCF Data management: filemaker server DB.	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Datasets containing non-personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Patients are coded as: CCSYRRRNNNN, 2 digits for city, 1 digit for Study. 1 digit for Year, 3 digit for Random letters. 4 digit for enrolment Number. Codes are pre generated for every enrolment kit.	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	Main data: own Vaccination registry	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	FISABIO	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:		Keywords or phrases describing the subjects or content of the data
Language:	Italian, English	All languages used in the dataset
Database size	filesize is 1.1 Mb	Indication of the size of the complete database
Technical description		
File inventory:	CIRI_Cohort18_Rev02.csv	All files associated with the project, including extensions
File formats:	csv	Format of the file
File structure:	According to DRIVE TND dataset requirements (ESSA)	Organization of the data file(s) and layout of the variables, where applicable
Access		

Rights	RGDP	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Data accessed CIRI-IT staff only	Where and how your data can be accessed by other researchers
Data sharing	No sharing. On demand with DRIVE	Description of how data will be shared, including access procedures
Ethics and legal issues	No problem	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	MUV	Name of the dataset or research project that produced it
Data owner:	Center for Virology, Medical University Vienna, Austria. Kinderspitalgasse 15, 1090 Vienna, Austria	Names and addresses of the organizations or people who own the data
Start and end date	Start date:4.10.18 End Date:13.5.19	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date:4.10.18 End date: 13.5.19	Start and end date of the period covered by the dataset
Methods:	primary data collection, Data obtained by face to face interview by the physicians Data management: Microsoft Access 2013, Microsoft Excel 2013	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Datasets containing age and gender, case control study based on data collected within the influenza surveillance system	Datasets containing personal data Datasets containing non-personal data
Processing:	The file was coded by the following procedure: identification numbers were deleted and each patient got an individual consecutive number, de-identification is not possible as it is fully anonymized	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	no data derived from other sources	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	Medical University Vienna, Sanofi Pasteur, Biomedica, Valneva	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:	case control study based on data collected within the influenza surveillance system	Keywords or phrases describing the subjects or content of the data
Language:	German	All languages used in the dataset
Database size	Xls-file: 253kb (ili) and 9kb (sari)	Indication of the size of the complete database
Technical description		

File inventory:	final_results_Austria_bis inkl. 14.04.2019_version2_ili.xls and final_results_Austria_bis inkl. 14.04.2019_version2_sari.xls	All files associated with the project, including extensions
File formats:	xls	Format of the file
File structure:	According to DRIVE's codebook	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights		Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Files are available upon request	Where and how your data can be accessed by other researchers
Data sharing	Aggregated, under the conditions of confidentiality required by Austrian law.	Description of how data will be shared, including access procedures
Ethics and legal issues	non	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	Case-control study to measure influenza vaccine effectiveness in Italy, 2018/19 season	Name of the dataset or research project that produced it
Data owner:	Istituto Superiore di Sanità Viale Regina Elena n. 2009, 00161 Rome, Italy	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 12/11/2018 End Date: 28/04/2019	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 15/10/2018 End date: 28/04/2019	Start and end date of the period covered by the dataset
Methods:	A test-negative case-control study was conducted within the context of the Influenza Surveillance Network (InfluNet). Data collection: electronic form on web-based platform (SQL Server) Data management: Stata 16.0 version	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Datasets do not contain personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Patients are coded as ID number	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	GPs collected data by an on-line standardized questionnaire, which converge in the Influenza Surveillance System database	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	Italian Medicines Agency	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:		Keywords or phrases describing the subjects or content of the data
Language:	Italian	All languages used in the dataset
Database size	400k (csv format)	Indication of the size of the complete database
Technical description		
File inventory:	DRIVE_Italy_2018_2019.csv	All files associated with the project, including extensions
File formats:	csv	Format of the file
File structure:	According to the DRIVE codebook	Organization of the data file(s) and layout of the variables, where applicable

Access		
Rights	RGDP	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Under request	Where and how your data can be accessed by other researchers
Data sharing	Aggregated data	Description of how data will be shared, including access procedures
Ethics and legal issues	No issues as the study has been approved by the Ethical Committee	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	DRIVE Project, National Institute for Infectious Diseases "Prof. Dr. Matei Balș", Bucharest, Romania	Name of the dataset or research project that produced it
Data owner:	National Institute for Infectious Diseases "Prof. Dr. Matei Balș" No. 1 Dr. Calistrat Grozovici street, Bucharest, 021105, Romania	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 12 November 2018 End Date: 30 April 2019	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 12 November 2018 End date: 30 April 2019	Start and end date of the period covered by the dataset
Methods:	Observational case-control study using the test-negative design. The study population consisted of patients seeking care (being admitted to hospital in the National Institute for Infectious Diseases "Prof. Dr. Matei Balș") with conditions related to influenza aged 6 months and above, with no contraindication for influenza vaccination. The study catchment area was defined as patients residing in Bucharest and the following geographical area: Ilfov, Dambovita, Giurgiu, Prahova, Arges, Teleorman, Ialomita, Dolj, Valcea, Olt. The denominator is the population of this geographical area (source population: 5 937 382 people). The study site is the National Institute for Infectious Diseases "Prof. Dr. Matei Balș", a reference center for infectious diseases in Romania, based in Bucharest, Romania. The National Institute of Infectious Diseases "Prof. Dr. Matei Bals" in Bucharest is the main center for infectious diseases	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)

	located in an urban area with a population of 2M+ inhabitants. The Institute has 740 beds in pediatrics and adults units (specifically 6 wards and 1 ICU for adult care and 4 wards and 1 ICU for pediatric care) equipped with the necessary appliances, at international standards, also certified ISO 9001/2000 for quality. The patients admitted in this center come not only from Bucharest but also from an area covering SE Romania with great impact on recruiting capabilities. The study was performed according to the DRIVE core protocol, separately from the influenza surveillance scheme, but data has been shared to and from other parallel influenza surveillance studies (e.g., Global Influenza Hospital Surveillance Network, local influenza surveillance in the Institute), as further described below.	
Type of data:	Dissociated personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Patients were allocated consecutive study numbers for de-identification.	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	Source: Own data of the National Institute for Infectious Diseases “Prof. Dr. Matei Balș”, collected for DRIVE and GIHSN (Global Influenza Hospital Surveillance Network).	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	DRIVE, National Institute for Infectious Diseases “Prof. Dr. Matei Balș” and GIHSN, which is funded by the Foundation for Influenza Epidemiology, France. The foundation is funded by Sanofi Pasteur.	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:	Influenza, vaccination, hospitalization, children, adults, elderly, comorbidities	Keywords or phrases describing the subjects or content of the data
Language:	English	All languages used in the dataset
Database size	253 kb	Indication of the size of the complete database
Technical description		
File inventory:	Baza de date DRIVE RO final v2 rev 21.05.2019.csv Study protocol DRIVE INBI Matei Bals v2 12.11.2018.docx	All files associated with the project, including extensions

	PO CC 01 - Obtinerea consimtamantului informat al subiectilor participanti la studii clinice.pdf	
File formats:	Csv, docx, pdf	Format of the file
File structure:	According to DRIVE codebook and core protocol	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	National Institute for Infectious Diseases “Prof. Dr. Matei Balș”	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Available from the principal investigator, upon reasonable request	Where and how your data can be accessed by other researchers
Data sharing	Aggregated, under the conditions of confidentiality required by Romanian law, only for scientific purposes as stipulated in the DRIVE agreement.	Description of how data will be shared, including access procedures
Ethics and legal issues	Ethical approval obtained from the Bioethics Committee of the National Institute for Infectious Diseases “Prof. Dr. Matei Balș”	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	Type/brand specific influenza vaccine effectiveness study	Name of the dataset or research project that produced it
Data owner:	Hospital Universitari Vall Hebron (Department of Preventive Medicine and Epidemiology), Passeig Vall Hebron 119-129 08035 Barcelona (Spain). This centre belongs to the ICS network (Institut Català de la Salut).	Names and addresses of the organizations or people who own the data
Start and end date	Start date: week 40 (October 2018) End Date: week 20 (May 2019)	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: December 13th 2018 End date: April 14th 2019	Start and end date of the period covered by the dataset
Methods:	A test-negative case-control study. Study participants (both cases and controls) were identified using the electronic clinical history of HUVH. Their clinical history was reviewed to confirm they comply with the case definition and the inclusion criteria. The data about the covariates was also obtained from their electronic clinical history. The information about the vaccination status was obtained from the vaccination record available in the primary healthcare clinical electronic record of Catalonia (ECAP) and/or the electronic clinical record of the hospital. Data was stored in a Microsoft Access® 2007 database and then exported to a Microsoft Excel®.CSV file.	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Dataset contains personal data (age, sex, clinical characteristics)	Datasets containing personal data Datasets containing non-personal data
Processing:	Data processing and cleaning was performed by the site investigators. The de-identification of the patients was carried out assigning to each patient an ID number based on the swab date and the clinical history number.	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	All data from the patients, including data collection of the	Citations to data derived from other sources, including

	covariates, was obtained from the hospital and primary healthcare clinical electronic records.	details of where the source data is held and how it was accessed
Funder:	DRIVE	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:		Keywords or phrases describing the subjects or content of the data
Language:	English	All languages used in the dataset
Database size	55 kB	Indication of the size of the complete database
Technical description		
File inventory:	VHUH DRIVE_ALL	All files associated with the project, including extensions
File formats:	CSV	Format of the file
File structure:	DRIVE codebook	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	GDPR	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Under request, after Ethics Committee Approval	Where and how your data can be accessed by other researchers
Data sharing	Aggregated data	Description of how data will be shared, including access procedures
Ethics and legal issues	None to our knowledge	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	Case control study to measure type/brand-specific seasonal influenza vaccine effectiveness against laboratory confirmed influenza hospitalisations in Italy, season 2018/19	Name of the dataset or research project that produced it
Data owner:	Single Hospitals collecting the data: • Liguria region: San Martino Polyclinic Hospital, Hygiene Unit. Department of Health Sciences, University of Genoa • Toscana region: AO Senese, Department of Molecular and Developmental Medicine. Lab Molecular Epidemiology University of Siena • Lazio region: Bambino Gesù Children's Hospital and Azienda Ospedaliero-Universitaria Sant'Andrea, Università degli Studi di Roma "La Sapienza" • Puglia region: AO Bari. Department of Biomedical Science, Università degli Studi di Bari	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 18/11/2018 End Date: 28/04/2019	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 20/11/2018 End date: 24/04/2019	Start and end date of the period covered by the dataset
Methods:	Primary data collection. Data obtained by face-to-face and telephone interview, by clinical records review and by contacting GPs for validating vaccination status. Prospective, active-surveillance, hospital-based study. TND. Data collection: electronic CRF on a web-based system (SQL) Data management: Stata 14.2	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)

Type of data:	Dataset containing non personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Patients are coded as: II-HH-NNNNNN, 2 digits for patient initial hospital, 2 digits for Hospital ID and 6 digits for patient onset date.	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	Main data. Vaccination information are collected through telephone interview with the GP	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	DRIVE	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:		Keywords or phrases describing the subjects or content of the data
Language:	Italian	All languages used in the dataset
Database size	1.7MB	Indication of the size of the complete database
Technical description		
File inventory:	IT_BIVE_HOSP.csv	All files associated with the project, including extensions
File formats:	csv	Format of the file
File structure:	According with the DRIVE Codebook	Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	RGDP	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Under request	Where and how your data can be accessed by other researchers
Data sharing	Aggregated data	Description of how data will be shared, including access procedures
Ethics and legal issues	No issues as the Study has been approved by the Ethical Committee	Description of any ethics and legal issues associated with the dataset, if any

DRIVE dataset details to be completed for each dataset

General Overview		Explanation
Title:	VAHNSI (Valencia Hospital Surveillance Network for the Study of Influenza) 2018/2019	Name of the dataset or research project that produced it
Data owner:	FISABIO – Public Health (Vaccine Research Department) Avenida Cataluña, 21, 46020 (València, Spain)	Names and addresses of the organizations or people who own the data
Start and end date	Start date: 08-10-2018 End Date: 05-06-2019	Study start and end date (dd-mm-yyyy)
Time period covered by the dataset:	Start date: 08-10-2018 End date: 12-05-2019	Start and end date of the period covered by the dataset
Methods:	Primary data collection. Data obtained by face-to-face interview, by clinical records review and by vaccination registries. Prospective, active-surveillance, hospital-based study. TND. Data collection: electronic CRF Data management: Stata MP 14.2	How the data were generated (e.g. primary data collection, registry, study design, etc.), listing equipment and software used (including model and version numbers)
Type of data:	Dissociated personal data	Datasets containing personal data Datasets containing non-personal data
Processing:	Patients are coded as: HH-NNNN, 2 digits for admission hospital and 4 digits for patient number (automatically given by the electronic CRF).	How the data have been altered or processed (e.g. normalized), including de-identification procedures
Source:	Main data: Own data Vaccination data: Valencia Region Vaccine Information System (VRVIS)	Citations to data derived from other sources, including details of where the source data is held and how it was accessed
Funder:	FISABIO – Public Health, CIBERESP (Instituto de Salud Carlos III) and Sanofi Pasteur	Organizations or agencies who funded the research, or indicate that the data owner funds the study
Content description		
Subject:		Keywords or phrases describing the subjects or content

		of the data
Language:	English	All languages used in the dataset
Database size	1.54MB (vahnsi_20190516.csv)	Indication of the size of the complete database
Technical description		
File inventory:	vahnsi_20190516.csv	All files associated with the project, including extensions
File formats:	csv	Format of the file
File structure:		Organization of the data file(s) and layout of the variables, where applicable
Access		
Rights	RGDP	Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data
Access information	Under request	Where and how your data can be accessed by other researchers
Data sharing	Aggregated, under the conditions of confidentiality required by Spanish law.	Description of how data will be shared, including access procedures
Ethics and legal issues		Description of any ethics and legal issues associated with the dataset, if any

Annex 6. Research Collaborator Agreement

RESEARCH COLLABORATOR AGREEMENT

(hereinafter the **"Agreement"**)

THIS AGREEMENT effective as of **[…]** (the "**Effective Date**"), by and between:

(1) [DRIVE PARTNER´S LONG NAME-SHORT NAME], validly represented by [NAME OF PERSON], [POSITION HELD AT THE ORGANIZATION], whose administrative offices are at [LEGAL ADDRESS];

(2) **FUNDACIÓN PARA EL FOMENTO DE INVESTIGACIÓN SANITARIA Y BIOMÉDICA DE LA COMUNITAT VALENCIANA [FISABIO]**, whose administrative offices are at C. Micer Mascó 31, 46010 Valencia (SPAIN)

Hereinafter referred to as individually, a "**Party**" and, collectively, the "**Parties**".

WHEREAS:

FISABIO has entered into, together with [Partner SHORT NAME] and various other legal entities (the "Participants"), a consortium agreement (the "Consortium Agreement") under the Innovative Medicines Initiative 2 ("IMI") for the Development of Robust and Innovative Vaccine Effectiveness (DRIVE) project, with Grant Agreement number 777363 (the "Project").

In any case of conflict between the regulations of this Research Collaborator Agreement and the Consortium Agreement, the Consortium Agreement shall prevail.

[Partner SHORT NAME] has been selected following the rules and eligibility criteria published in the DRIVE *Call for tenders -2018/2019 influenza season* to extend the DRIVE network by including new study sites capable of producing Influenza Vaccine Effectiveness (IVE) estimates. In this sense, the Description of the Work to be carried out by [Partner SHORT NAME] and provided to DRIVE is defined in **Annex 1** of this Agreement ("the "**Work**").

[Partner SHORT NAME] agrees, on the conditions set forth below, to participate to the Project performing the Work as identified in Annex 1 under the terms and conditions as defined herein.

NOW, THEREFORE, IT IS AGREED AS FOLLOWS:

1. DEFINITIONS

1.1. The following capitalized terms used in this Agreement shall have the following meanings:

a. "**Beneficiary**" or "**Participant**" means a legal entity who has signed Grant Agreement 777363 with the IMI2 JU or the Form of Accession. The current

list of beneficiaries is included in **Annex 2.**

b. "**Confidential Information**" means any data, documents or other material (in any form) that is identified as confidential in writing and marked as confidential (or with other similar designation) at the time of disclosure.
c. "**Description of Work**" means all the activities to be performed by [Partner SHORT NAME] to contribute to the Project which are defined in the Annex 1 of this Agreement.
d. "**Data**" means any and all data and all supporting information, related reports, any and all results, whether patentable or not, arising directly or indirectly, in the performance of the Work under this Agreement, including mainly the data collected by [Partner SHORT NAME]. The definition of Data can also encounter existing data prior collected by the beneficiary that could be useful for the Work performance.
e. "**Personal Data**" means any information relating to an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his/her physical, physiological, mental, economic, cultural or social identity.

2. CONTRIBUTION TO THE PROJECT

2.1. DATA CONTRIBUTION

2.1.1. The Data collected by [Partner SHORT NAME] for DRIVE specific needs/objectives will be provided to the Participant named "P95 CVBA (P95)", the DRIVE partner responsible for the pooled/meta-analysis, located in Belgium. [Partner SHORT NAME] will remain the owner of the Data as more explained in Clause 4. The Data generated by [Partner SHORT NAME] will be accessible only to the non-EFPIA partners of DRIVE for scientific review and pooled/meta-analysis purposes (i.e. no access by the pharmaceutical companies) and, if necessary, to a third party Contract Research Organization (CRO) commissioned by DRIVE's Quality Control and Audit Committee for auditing purposes.

2.1.2. The Data will be transferred to P95 via secure file transfer protocol as follows: a representative of [Partner SHORT NAME] that has been granted access to the DRIVE Research Server, transfers the Data files to the server using an application that allows for Data transfer over the secure file transfer protocol. This allows the Data to be transferred securely and encrypted. Upon arrival of the Data the server admin will perform the necessary safety and compliance checks.

2.1.3. [Partner SHORT NAME] confirms that all applicable consents and approvals are in place for the Participants and [Partner SHORT NAME] to undertake their respective activities under Clause 3.

2.1.4. [Partner SHORT NAME] may transfer the Data to other Participants, Affiliates and/or Third Parties for purposes of completion of the Project or for Research Use as defined in Clause 4 of this Agreement. The Participants, Affiliates and/or Third Parties receiving the Data will be required to abide by the same obligations as provided for in this Agreement. The Data will under no circumstances be transferred

to vaccine manufacturing authorization holders.

2.1.5. The Parties agree that the Participants can request more detailed (but fully anonymized) information on the Data generated in the activities described in Annex 1. [Partner SHORT NAME] shall undertake reasonable efforts to respond thereto within a reasonable timeframe.

2.1.6. [Partner SHORT NAME] agrees that the responses to the queries can be shared between the Participants as well as with their Affiliates and Third Parties to the extent such sharing takes place in accordance with the terms and conditions of all applicable legal terms governing the Project (mainly the Grant and Consortium Agreements).

2.1.7. [Partner SHORT NAME] will always provide Personal Data duly anonymized or codified so that the DRIVE consortium will not be able to identify any physical subject.

2.1.8. [Partner SHORT NAME] must process any Personal Data under this Agreement in compliance with applicable EU General Data Protection Regulation (EU) 2016/679 ("GDPR") and national laws on data protection (including, without being limited to, registration, authorisation or notification requirements). [Partner SHORT NAME] represents and warrants that (i) any Personal Data required to perform the Work are collected, obtained, handled, hosted, transferred or used by it or its Affiliated Entities, Linked Third Parties, or Sub-Contractors will be obtained, handled or used in accordance with all relevant laws and regulations (and where applicable, local ethical guidelines) (ii) it will, and will ensure that its Affiliated Entities, Linked Third Parties or Sub-Contractors, will adhere to the principles of medical confidentiality in relation to any patient involved in studies conducted as a part of the Work; (iii) when sponsoring a study forming part of the Work, secure the written consent of each patient prior to using, collecting, disclosing or transferring any Personal Data of that patient (including in particular its identity) to any third party, except as permitted by applicable EU and national laws on data protection; and (iv) that any Ethics Committee approvals and physical subjects informed consents required for performing the Work will be obtained prior to the commencement of the respective actions. [Partner SHORT NAME] may grant their personnel access only to such Personal Data if this is strictly necessary for implementing, performing, managing and monitoring the Work.

3. COMPENSATION

3.1. [Partner SHORT NAME] shall be compensated through a budget reallocation from FISABIO (the DRIVE Coordinator) budget devoted to studies to [Partner SHORT NAME] budget for its actual effort in the Project for the development of the Work described in **Annex 1**.

3.2. [Partner SHORT NAME] shall be compensated by FISABIO for the **costs actually incurred** by [Partner SHORT NAME] (i.e. personnel, equipment depreciation, consumables, travel and subsistence allowances and 25% maximum of indirect

costs). Eligibility conditions of incurred costs are included in **Annex 3** of this Agreement.

3.3. The allocated budget will be appropriately sized to the related Work. The maximum budget to be transferred by FISABIO is included **Annex 4** of this Agreement; double funding (the situation where the same activity would be paid twice from different funding sources) will not be possible.

3.4. Any transfer of funding will be made by FISABIO within thirty (30) days after the achievement of the following milestones:

1. End of January 2019: reception and acceptance of the study protocol (based on a DRIVE generic protocol). Transfer of a pre-financing of 30% of the estimated budget (Annex 4) **[include amount]**
2. End of September 2019: reception and acceptance of the Study Report and the Financial Reporting Form (Annex 5). The rest of funding until 100% of the declared expenses with a maximum of the approved budget **[include amount]**

3.5. The Coordinator will do the funding transfer to a bank account nominated by [Partner SHORT NAME] in clause 3.7 and complying with all applicable legal and tax requirements. [Partner SHORT NAME] acknowledges and agrees that the transferred amounts will be reported to the Partners of the Consortium.

3.6. [Partner SHORT NAME] shall be responsible for all other taxes payable on account of payments made hereunder.

3.7. The grant must not produce a profit. "Profit" means the surplus of the amount obtained over the action's total eligible costs. [Partner SHORT NAME] shall not receive any unconnected benefits from the execution of this contract beyond the compensation agreed herein.

3.8. All funding transfers made by FISABIO shall be made into the following account of [Partner SHORT NAME]:

Account holder: _____
Bank: _____
Bank address:_____
Bank identification number (if applicable):_____
Account number:_____
IBAN:_____
SWIFT/BIC CODE:_____
Reason for payment: _____

4. ACCESS RIGHTS

4.1. The Data owned before execution of this Agreement or developed by [Partner SHORT NAME] during this Agreement validity shall remain the sole property of the [Partner SHORT NAME].

4.2. [Partner SHORT NAME] will be free to publish its own Data. DRIVE funding for primary Data collection shall be acknowledged as per ICJME guidelines and DRIVE may receive the publication for non-binding comments.

4.3. [Partner SHORT NAME] hereby grants access rights to the Project Participants to its Data, without any additional payment than the compensation agreed herein, for the implementation of the Project in accordance with the Consortium Agreement terms.

4.4. During the term of this Agreement and for such additional period of time that records are required to be retained by law or otherwise, it is agreed that authorised external auditors may arrange audit with [Partner SHORT NAME], at regular business hours of the locations where the Data collection is performed and after having duly informed [Partner SHORT NAME] respecting at least seven (7) days prior written notice, in order to examine and audit the locations where the Data collection is performed and to, subject to confidentiality considerations, inspect, audit, copy or have copied, all data and document generated relating to the Work performed under this Agreement in order to confirm that Data collection is performed in accordance with the Protocol and all applicable laws and regulatory requirements

[Partner SHORT NAME] agrees to assist the auditor to the extent deemed reasonable, in facilitating examination, inspection, auditing and copying of materials relating to the Work. [Partner SHORT NAME] agrees to take any action, as reasonably requested by FISABIO, to properly correct or address any deficiencies noted during any audit.

The external auditors will maintain the confidentiality of any subject-identifiable medical records.

5. REPRESENTATIONS AND WARRANTIES

5.1 FISABIO:

5.1.1 ensures that the Data will be stored in a secure and appropriate manner and in accordance with all applicable laws and regulations.

5.1.2 undertakes to comply with all applicable laws, rules, regulations, ordinances, and directives.

5.2 [Partner SHORT NAME]:

5.2.1 undertakes to comply with all applicable laws, rules, regulations, ordinances, and directives.

5.2.2 represents and warrants that all necessary consents have been obtained for the Data to be used in the Project and for Research Use as defined in Clause 4.

5.2.3 represents and warrants that the Personal Data provided is duly anonymized or codified.

5.2.4 ensures that IMI, the European Commission, the European Court of Auditors (ECA) and the European Anti-Fraud Office (OLAF) can exercise their rights of Checks, Reviews and Audits as well as Evaluation of the Impact if necessary as established in the Project Grant Agreement.

5.3 Each Party represents and warrants that (i) it has the full right and authority to enter into and perform its obligations as set forth in this Agreement; (ii) the execution, delivery and performance of this Agreement have been duly authorized by all necessary corporate action on its part; (iii) it will not grant any rights in conflict with any of the rights granted hereunder; (iv) the rights granted pursuant to this Agreement do not and will not infringe with any Third Party rights and, notwithstanding the foregoing, it will notify the other Parties promptly after becoming aware of any Third Party right infringement.

6. TERM AND TERMINATION

6.1. Unless terminated earlier in accordance with the provisions of this Section 8, the term of this Agreement shall commence on **[…]** (the Effective Date) and shall remain in effect **until the fulfillment of respective obligations of both Parties** (after Work performed by [Partner SHORT NAME] and last payment done by FISABIO).

6.2. Termination for Breach: either party may terminate this Agreement if the other party materially breaches a provision of this Agreement and fails to cure such breach within thirty (30) days of receipt of written notice describing the breach in reasonable detail.

7. MISCELLANEOUS

7.1. This Agreement shall be governed by and interpreted in accordance with the law of Spain excluding the application of its conflict laws.

7.2. Any dispute shall be finally settled by arbitration in Valencia, Spain under the rules of arbitration of the Ilustre Colegio Oficial de Abogados de Valencia.

THIS AGREEMENT has been executed in two (2) originals to be duly signed by the undersigned authorised representatives.

[Partner SHORT NAME]	**[FISABIO]**
By: _____	By: _____
Position: _____	Position: _____
Date: _____	Date: _____

ANNEX 1

DESCRIPTION OF THE WORK TO BE PROVIDED BY THE RESEARCH COLLABORATOR

ANNEX 2
DRIVE PROJECT BENEFICIARIES

1. FUNDACION PARA EL FOMENTO DE LA INVESTIGACION SANITARIA Y BIOMEDICA DE LA COMUNITAT VALENCIANA (FISABIO), established in CALLE MICER MASCO 31, VALENCIA 46010, Spain

2. INSTITUT DE RECHERCHE POUR LE DEVELOPPEMENT (IRD), established in BOULEVARD DE DUNKERQUE 44 CS 90009, MARSEILLE 13572, France

3. P95 CVBA (P95), established in KONING LEOPOLD III LAAN 1, HERVELEE 3001, Belgium

4. UNIVERSITA DEGLI STUDI DI FIRENZE (UNIFI), established in Piazza San Marco 4, Florence 50121, Italy

5. SYNAPSE RESEARCH MANAGEMENT PARTNERS SL (SYNAPSE), established in CALLE DIPUTACION 237 PLANTA AT PUERTA 3, BARCELONA 08007, Spain

6. TERVEYDEN JA HYVINVOINNIN LAITOS (THL), established in MANNERHEIMINTIE 166, HELSINKI 00271, Finland

7. ISTITUTO SUPERIORE DI SANITA (ISS), established in Viale Regina Elena 299, ROME 00161, Italy

8. UNIVERSITY OF SURREY (SURREY), established in Stag Hill, GUILDFORD GU2 7XH, United Kingdom

9. CONFEDERATION OF MENINGITIS ORGANISATIONS LTD (COMO), established in NEWMINSTER HOUSE BALDWIN STREET, BRISTOL BS1 1LT, United Kingdom

10. UNIVERSITE LYON 1 CLAUDE BERNARD (UCBL), established in BOULEVARD DU 11 NOVEMBRE 1918 NUM43, VILLEURBANNE CEDEX 69622, France

11. ASSOCIATION INTERNATIONALE DE STANDARDISATION BIOLOGIQUE POUR L'EUROPE (IABS-EU) (IABS-EU), established in BOULEVARTD VIVIER MERLE 10-12 WORLD TRADE CENTER TOUR OXYGENE, LYON 69393, France

12. SANOFI PASTEUR SA (SP), established in 14, Espace Henry Vallée - 69007 LYON – France

13. ABBOTT BIOLOGICALS BV (ABBV), established in C.J. VAN HOUTENLAAN 36, WEESP 1381 CP, Netherlands

14. SEQIRUS UK LIMITED (SEQIRUS), established in POINT 29 MARKET STREET, MAIDENHEAD SL6 8AA, United Kingdom

15. GLAXOSMITHKLINE BIOLOGICALS (GSK Bio), established in Rue de l'Institut 89, Rixensart 1330, Belgium

ANNEX 3
CONDITIONS FOR ELIGIBILITY OF COSTS

[Partner SHORT NAME] must have the appropriate resources to implement the action. If it is necessary to implement the action related to DRIVE, they may purchase goods, works and services. [Partner SHORT NAME] retains sole responsibility towards DRIVE consortium for implementing the action. The **grant must not produce a profit**. 'Profit' means the surplus of the amount obtained over the action's total eligible costs.

ELIGIBILITY OF COSTS

1. FORM OF GRANT, REIMBURSEMENT RATES AND FORMS OF COSTS

The grant reimburses 100% of the action's eligible costs.
Eligible costs must be declared under the following forms ('forms of costs'):

a) for direct personnel costs:
b) for direct costs for subcontracting: as actually incurred costs (actual costs);
c) for other direct costs: as actually incurred costs (actual costs);
d) for indirect costs: on the basis of a 25% flat-rate

2. GENERAL CONDITIONS FOR COSTS TO BE ELIGIBLE

'**Eligible costs**' are costs that meet the following criteria:

i. they must be actually incurred by the beneficiary;
ii. they must be incurred in the period set out in proposal;
iii. they must be indicated in the estimated budget;
iv. they must be incurred in connection with the action as described in the proposal and necessary for its implementation;
v. they must be identifiable and verifiable, in particular recorded in the beneficiary's accounts in accordance with the accounting standards applicable in the country where the beneficiary is established and with the beneficiary's usual cost accounting practices;
vi. they must comply with the applicable national law on taxes, labour and social security;
vii. they must be reasonable, justified and must comply with the principle of sound financial management, in particular regarding economy and efficiency;

2.A Direct personnel costs

Personnel costs are eligible, if they are related to personnel working for the beneficiary under an employment contract (or equivalent appointing act) and assigned to the action. They must be limited to salaries (including during parental leave), social security contributions, taxes and other costs included in the remuneration, if they arise from national law or the employment contract (or equivalent appointing act).

Personnel costs = hourly rate x number of actual hours worked on the action

The number of actual hours declared for a person must be identifiable and verifiable. The total number of hours declared for a person for a year, cannot be higher than the annual productive hours used for the calculations of the hourly rate.

For personnel costs declared as the hourly rate is the amount calculated as follows:

Hourly rate = actual annual personnel costs for the person / number of annual productive hours

For the '**number of annual productive hours'**, the beneficiaries may choose one of the following:

i. 'fixed number of hours': 1.720 hours for persons working full time (or corresponding pro-rata for persons not working full time);

ii. 'individual annual productive hours': the total number of hours worked by the person in the year for the beneficiary, calculated as follows:

{annual workable hours of the person (according to the employment contract, applicable collective labour agreement or national law) + overtime worked - absences (such as sick leave and special leave)}.

'Annual workable hours' means the period during which the personnel must be working, at the employer's disposal and carrying out his/her activity or duties under the employment contract, applicable collective labour agreement or national working time legislation. If the contract (or applicable collective labour agreement or national working time legislation) does not allow determining the annual workable hours, this option cannot be used;

iii. 'standard annual productive hours': the standard number of annual hours generally applied by the beneficiary for its personnel in accordance with its usual cost accounting practices. This number must be at least 90% of the 'standard annual workable hours'.

If there is no applicable reference for the standard annual workable hours, this option cannot be used.

For all options, the actual time spent on parental leave by a person assigned to the action may be deducted from the number of annual productive hours.

2.B Direct costs of subcontracting

If necessary to implement the action, the beneficiaries may award subcontracts covering the implementation of certain action tasks described in the proposal. Subcontracting may cover only a limited part of the action. The beneficiaries must award the subcontracts ensuring the best value for money or, if appropriate, the lowest price. In doing so, they must avoid any conflict of interests. The tasks to be implemented and the estimated cost for each subcontract must be set out in the proposal.

2.C Other direct costs

C.1 Travel costs and related subsistence allowances (including related duties, taxes and charges such as non-deductible value added tax (VAT) paid by the beneficiary) are eligible if they are in line with the beneficiary's usual practices on travel.

C.2 The depreciation costs of equipment, infrastructure or other assets (new or second-hand) as recorded in the beneficiary's accounts are eligible and written off in accordance with international accounting standards and the beneficiary's usual accounting practices.
The **costs of renting or leasing** equipment, infrastructure or other assets (including elated duties, taxes and charges such as non-deductible value added tax (VAT) paid by the beneficiary) are also eligible, if they do not exceed the depreciation costs of similar equipment, infrastructure or assets and do not include any financing fees.
The costs of equipment, infrastructure or other assets **contributed in-kind against payment** are eligible, if they do not exceed the depreciation costs of similar equipment, infrastructure or assets and do not include any financing fees.
The only portion of the costs that will be taken into account is that which corresponds to the duration of the action and rate of actual use for the purposes of the action.

C.3 Costs of other goods and services (including related duties, taxes and charges such as non-deductible value added tax (VAT) paid by the beneficiary) are eligible, if they are:
(a) purchased specifically for the action or
(b) contributed in kind against payment
Such goods and services include, for instance, consumables and supplies, dissemination (including open access), protection of results, translations and publications.

2. D. Indirect costs
Indirect costs are eligible if they are declared on the basis of the flat-rate of 25% of the eligible direct costs from which are excluded:
(a) costs of subcontracting and
(b) costs of in-kind contributions provided by third parties which are not used on the beneficiary's premises;

3. INELIGIBLE COSTS
'**Ineligible costs**' are:

1) costs related to return on capital;
2) debt and debt service charges;
3) provisions for future losses or debts;
4) interest owed;
5) doubtful debts;
6) currency exchange losses;
7) bank costs charged by the beneficiary's bank for transfers from the JU;
8) excessive or reckless expenditure;
9) deductible VAT;
10) costs incurred during suspension of the implementation of the action;
11) costs declared under another JU, EU or Euratom grant.

4. CONSEQUENCES OF DECLARATION OF INELIGIBLE COSTS
Declared costs that are ineligible will be rejected.

ANNEX 4
ESTIMATED BUDGET

[TO BE FILLED BY THE ASSOCIATED PARTNER ACCORDING TO THE BUDGET APPROVED IN THE CALL FOR TENDERS]

		Person-Months	Budget (€)	Description
A. Personnel costs (€)				
B. Other Direct costs	B.1 Travel and subsistence (€)			
	B.2 Consumables (€)			
	B.3 Other goods and services (€)			
C. Subcontracting (€)				
Total Direct Costs (€)				
F. Indirect costs (€)	(max. 25% Total Direct Costs except Subcontracting)			
	TOTAL BUDGET (€)			

ANNEX 5
FINANCIAL REPORTING FORM

(Only as reference, the original file will be provided by FISABIO)

Associated Partner [insert Entity name]			
Direct personnel costs declared as actual costs			
Persons months per WP:			
Person months *(1 pm is the measure of time of 1 person working full time within a month (depending on the organization, 1 pm usually varies between 140 h-160 h). To calculate the pm please refer to slide 14 of the Financial Seminar slides)*	*Please indicate the name of the personnel involved, hourly rate and number of hours dedicated to the project*	*Explanation of the activities carried out and specific achievements reached by your institution*	*Amount in EUR*
[insert number pm]		*(e.g authorship of the deliverable XX)*	[insert amount in EUR]
[insert number pm]		[insert comment]	[insert amount in EUR]
[insert number pm]		[insert comment]	[insert amount in EUR]
[insert number pm]		[insert comment]	[insert amount in EUR]
[insert number pm]		[insert comment]	[insert amount in EUR]
[insert number pm]		[insert comment]	[insert amount in EUR]
0			0

Direct costs of subcontracting - *if applicable*		
Costs	**Explanation**	**Foreseen in Annex 1 (DoA)**
0	[insert comment]	*[* YES*]* *[* NO*]*

Other direct costs				
Category	**Costs**	**Short Description**	**Associated WP**	**Explanation (if not included in Annex 1-DoA-)**
Travels:	[insert amount in EUR]	[insert date and location, name of the meeting, name of the attendees] *(e.g. PI (full name) attending DRIVE meeting held in Valencia, 17-19 July 2018)*	[insert WP number]	[insert comment]
	[insert amount in EUR]	[insert date and location, name of the meeting, name of the attendees]	[insert WP number]	[insert comment]
Other goods or services:	[insert amount in EUR]	[detail name and category. Please refer to the IMI2 Annotated Model Grant Agreement for further information]	[insert WP number]	[insert comment]
TOTAL	0			

TOTAL Costs claimed	**0,00 €**
Please include the total amount claimed as part of communication/dissemination activities:	